XXV. NEUROPHYSIOLOGY

Academic and Research Staff

Prof. Jerome Y. Lettvin
Prof. Stephen G. Waxman
Dr. Isabelle Alter
Dr. Michael H. Brill
Dr. Edward R. Gruberg
Dr. Eric A. Newman
Dr. George M. Plotkin
Dr. Stephen A. Raymond

Graduate Students

Larry R. Carley
Charles L. Epstein
James F. Green
Bradford Howland
Lynette L. Linden
Kenneth J. McLeod
Louis L. Odette
William M. Saidel

1. A CENTRAL PATHWAY OF THE INFRARED SYSTEM OF THE RATTLESNAKE, Crotalus viridis

National Institutes of Health (Training Grant 5 TOl EY00090)

Bell Laboratories (Grant)

Edward R. Gruberg, Eric A. Newman

[Eric A. Newman is with the Eye Research Institute, Retina Foundation, Boston, Mass.]

We have traced a pathway of the infrared system in the rattlesnake from the nucleus of the lateral descending trigeminal tract (LTTD, the primary infrared nucleus) to the tectum. This pathway links the LTTD and the intermediate tectal neuropil, where previous electrophysiological studies have demonstrated infrared responses. Following HRP injections into the intermediate layers of the tectum, the Mesulam benzidine blue method revealed a group of large (25-40 μ) cells on the ventrolateral margin of the contralateral medulla, filled by retrograde transport. This nucleus, which is distinct from the primary trigeminal nuclei, we have provisionally named "nucleus R". Its rostral boundary is immediately posterior of the Vth root entry, and its caudal boundary is posterior to the point of closure of the fourth ventricle. The axons from the nucleus R decussate in the ventral medulla, turn rostral in a ventral bundle, and in the tegmentum proceed dorsolaterally to the tectum. Extracellular microelectrode recordings made from the region of the nucleus R show that units are driven by infrared but not by visual or tactile stimuli. Electrolytic lesions made by the recording electrodes confirm that the units are in the nucleus R. Following HRP injection into the nucleus R, cells in the ipsilateral LTTD are heavily stained due to retrograde transport. The axons of nucleus R cells are also stained in these experiments (by anterograde transport), with terminal branches in intermediate tectal layers. The axons follow the same nucleus R-tectal pathway seen following tectal HRP injections. These experiments show that an infrared pathway in the rattlesnake proceeds from the LTTD to the ipsilateral nucleus R and thence to the contralateral tectum.

2. CONNECTIONS OF THE TECTUM OF THE RATTLESNAKE, Crotalus viridis: AN HRP STUDY

National Institutes of Health (Training Grant 5 TOl EY00090)

Bell Laboratories (Grant)

Edward R. Gruberg, Eric A. Newman

[Eric A. Newman is with the Eye Research Institute, Retina Foundation, Boston, Mass.]

This research has been done in collaboration with E. Kicliter, Department of Anatomy and Laboratory of Neurobiology, University of Puerto Rico, and P. H. Hartline and L. Kass, Eye Research Institute, Retina Foundation, Boston, Mass.

We have studied the connections of the tectum of the rattlesnake by tectal application of horseradish peroxidase. The tectum receives bilateral input from nucleus lentiformis mesencephali, posterolateral tegmental nuclei, anterior tegmental nuclei, and periventricular nuclei; ipsilateral input from nucleus geniculatus pretectalis, and lateral geniculate nucleus pars dorsalis; and contralateral input from dorsolateral posterior tegmental nucleus and the previously undescribed nucleus reticularis caloris (RC). RC is located on the ventrolateral surface of the medulla and consists of large cells 25-40 μ in diameter. Efferent projections from the tectum can be traced to the ipsilateral nucleus lentiformis mesencephali, the ipsilateral lateral geniculate region, anterior tegmental region, and wide bilateral area of the neuropil of the ventral tegmentum and ventral medulla. We have not found any direct tectal projections from the sensory trigeminal nuclei including the nucleus of the lateral descending trigeminal tract (LTTD). We suggest that in the rattlesnake RC is the intermediate link connecting LTTD to the tectum.

3. VOLUMETRIC THEORY OF COLOR CONSTANCY

Bell Laboratories (Grant)

Michael H. Brill

In previous papers[1,2] we presented a design for a trichromatic photosensing device with an analog of color constancy: For a particular class of illuminants, object spectral reflectances, and spectral sensitivities (tristimulus functions in a particular basis) of the device, we proved the illuminant invariance of any ratio of tristimulus volumes generated by triads of object colors. In that work, it was assumed that the illuminant energy spectrum is a linear combination of three known functions of wavelength.

We recently showed[3] that a single, different assumption is sufficient to assure the

illuminant-invariance of tristimulus volume ratios – it is enough that object spectral reflectances be linear combinations of three known functions of wavelength.

Reflectance spectra for natural, nonmetallic objects (see Ref. 4) are smoothly varying and generally have, at most, two maxima in the human visible-wavelength range; this suggests that such spectra may be reasonably represented by an expansion in terms of three basis functions over this range.

Although the present assumption constrains the object reflectance spectra more heavily than in the previous paper, it allows almost complete freedom for the possible illuminants. For example, it gives illuminant invariance when tungsten lights are replaced by the prime-color lights developed by W. A. Thornton (Ref. 5), which have just three spectral lines but are marketed for their high efficiency and color-rendering capability. (One must insure only that the illuminant does not render all object colors coplanar in tristimulus space, as it would if it consisted of one or two spectral lines – or more generally, if det [A] = 0 in the treatment that follows; in such cases, a volume ratio cannot be defined.) Thus the present invariance occurs under conditions like those under which color constancy obtains for human vision – a feature that vindicates and transcends our original intention of designing an illuminant-invariant object-color recognizer.

The mathematical invariance is readily shown as follows: Let the device's tristimulus functions be $q_j(\lambda)$ (j = 1, 2, 3), the illuminant energy spectrum be $I(\lambda)$, and the reflectance spectrum of the i^{th} object be $r_i(\lambda) = \sum_{k=1}^{3} B_{ik} u_k(\lambda)$. Here, it is assumed that all reflectance spectra in nature differ only in their B-parameters. Then the linear filtrate (tristimulus value) for the i^{th} object due to the j^{th} tristimulus function is

$$Q_{ij} = \int I(\lambda)\, r_i(\lambda)\, q_j(\lambda)\, d\lambda = \sum_{k=1}^{3} B_{ik} \int I(\lambda)\, q_j(\lambda)\, u_k(\lambda)\, d\lambda \equiv \sum_{k=1}^{3} B_{ik} A_{kj}.$$

Given three objects i, this equation can be written in the square-matrix form [Q] = [B][A]. Here [B] depends on the reflectance spectra but not on the illuminant. [A] depends on the illuminant, and on the tristimulus and reflectance-basis functions, but not on the reflectance parameters.

The volume of the parallelepiped in tristimulus space generated by these three objects is det [Q] = det [A] det [B]. Another parallelepiped from three other objects under the same light will have a volume given by det [Q'] = det [A] det [B']. The ratio of these volumes is det [Q]/det [Q'] = det [B]/det [B'], which is manifestly illuminant-invariant.

Previously[1, 2] we restricted the illuminant spectrum to be an expansion in three basis functions; this did not free the reflectance spectra to the extent the reciprocal arrangement did for the illuminant spectrum in the present formation. Clearly,

reflectance and illuminant spectra do not play symmetric roles in the volumetric approach to color constancy. The present assumption solves the problem of color constancy when looking through a nonturbid colored medium of variable thickness, which is important, for example, for the survival of underwater creatures and fishing land-dwellers. In the previous work, we assumed that the illuminant spectrum is a linear combination of three known functions, and is thereby defined as "smoothly varying." If two spectra satisfy this assumption, however, in general their product will not. Thus varying the thickness of water through which one looks must cause departures from the stated assumption of those papers.

With the present assumption, an otter fishing in varying depths of water could get depth-independent reflectance information from tristimulus volume ratios – so long as the reflectance spectra in its environment were well-behaved in the sense described above. Even through turbid water, reflectance information can be obtained from ratios of volumes produced by tristimulus-vector differences; the differencing operation removes dependence on added light reflected directly from the turbid medium.

References

1. M.H. Brill, J. Theor. Biol. 71, 473 (1978).
2. M.H. Brill, M.I.T. RLE Progress Report No. 120, January 1978, pp. 169-171.
3. M.H. Brill, "Further Features of the Illuminant-Invariant Trichromatic Photosensor," J. Theor. Biol. 78, 305 (1979).
4. G. Wyszecki and W. S. Stiles, Color Science (John Wiley and Sons, 1967), pp. 179-187.
5. W.A. Thornton, Light. Des. Appl. (November 1975), p. 35.

4. PERCEPTION AND THE OBSERVER IN RELATIVITY THEORY

Bell Laboratories (Grant)

Michael H. Brill

If only by analogy, the study of perception can clarify issues in physics. Sometimes we can use perceptual considerations to find common-experience analogs of "counter-intuitive" physical propositions. For example, consider the following proposition in Special Relativity: If A and B are co-moving observers, and each holds a ruler in the direction of their relative velocity, then each sees the other's ruler as shorter than his own. This proposition has a more mundane perceptual analog: A and B stand twenty feet apart and each holds a ruler vertically at arm's length. Then A's ruler subtends a larger visual angle to A than B's ruler, and B's ruler subtends a larger visual angle to B than A's ruler. This example shows how observed relationships between objects

can become symmetric when observer labels change together with object labels.

Not all the lessons of perception render established physical theories intuitively plausible. One lesson in particular, that different states of the world – such as metameric colors – can be indistinguishable to an observer, might cause one to look again at some seemingly incontrovertible connections between theory and experiment. Consider the famous Michelson-Morley experiment, which showed that the round-trip travel time of light is isotropic (the same in all directions). Does this imply that the speed of light is isotropic? It is easily shown that the following eikonal (direction dependence of the speed of light) gives the same round-trip travel times for all directions, even though it prescribes a nonisotropic speed of light:

$$u(\theta,\phi) = \frac{c}{1 + \epsilon \cos(\theta-\theta_o)},$$

where u is the speed of light, θ and ϕ are angular coordinates, and c, ϵ, and θ_o are constants.

This is an ellipsoid with two equal axes and the observer at one focus. It is only one – the simplest – eikonal that does the trick. It is implicit in an alternative to special-relativistic kinematics based on clock synchronization other than by electromagnetic waves.[1-3]

The above example does not challenge Special Relativity (which has many more empirical confirmations than the Michelson-Morley experiment), but illustrates the caution that must be taken in connecting measurement – and the observer – with physical theory. The numbers that emerge from a theory reflect implicit assumptions about the observer.

Let us examine this point as it relates to General Relativity, which promises deference to the observer via the Principle of Equivalence and the Principle of General Covariance. Most liberally interpreted, the Principle of Equivalence states that the local effects of a gravity field near an observer are indistinguishable from the effects of accelerating the observer,[4] and the Principle of General Covariance states that the observations of the universe by different observers can be connected by a coordinate transformation.[5]

Although both these principles guided Einstein's formulation of the gravitational field equations, no prescription has been derived from the field equations for determining the change in the observed universe when the observer undergoes acceleration. By itself, this is not a problem: The coordinate-independent formulation of the field equations cannot be expected to carry the observer dependence if this dependence resides in choice of coordinates (as dictated by the Principle of General Covariance). Solving the field equations, however, involves selecting a coordinate condition (boundary condition) on the equations. If observer frames of reference can be connected by a coordinate transformation, it is in the coordinate condition that we would expect to find the

prescription of observer dependence. Such a prescription has not been incorporated in solutions of the field equations in which there are gravitational sources (i. e., in which the space is not flat). Eddington's idea of "proper coordinates"[6] is an acknowledgment of the problem, but without a detailed formulation. The field equations with gravitational sources have been solved for only a few special cases, none of which parametrically incorporates the state of motion of the observer.

These considerations bear significantly on the empirical interpretation of nonflat solutions to Einstein's field equations. The fundamental tests of General Relativity – light-bending, red shift, and the perihelion advance of mercury – all invoke the static, spherically symmetric Schwarzschild solution, which has a coordinate condition at spatial infinity. It might be argued that this coordinate condition is a statement about the universe seen by an observer on earth, but then how would this vary when the observer is not on the earth? If General Relativity is to be able to distinguish between observers on and off the earth in its description of such effects as the red shift of light from the sun (which is certainly influenced by observer velocity, for example), it must include in its explanation of present earth-bound observations the fact that the earth is neither at spatial infinity nor static with respect to the sun. To do this, the theory must be augmented to incorporate the observer's state of motion and position in the gravity field.

One possible approach is to extend the approach of Anderson[7] to nonflat spaces. By a coordinate transformation, Anderson extended Special Relativity to a uniformly accelerated observer in a gravitational-source-free space: Given the universe seen by inertial observer A, accelerating observer B's world line in that description, and the acceleration sensed by B, B's description of the universe can be inferred. Ultimately, the acceleration sensed by B should prove to be unnecessary information if the Principle of Relativity is to be suitably generalized – i. e., if any observer's picture of the universe is to be complete. If an observer's picture of the universe is incomplete (as it is, for example, in visual perception when one viewed object eclipses another in depth), an additional "observer-free" description is required to make a theory predictive. Such an "objective" description of the universe is foreign to the ideas of Relativity.

H. Yilmaz[8] has developed an alternative to Einstein's theory with static solutions in which the observer's position in the gravity field figures explicitly. We are now exploring more general observer dependences in the theory, within which a change of observer is not simply a coordinate transformation. We hope to be able to predict such things as the appearance of the perihelion advance of Mercury to an observer in space co-moving with respect to the earth. Any claims that the effects of such observer changes are small must be supported by approximations within the gravity theory being tested (not from Special Relativity or Newtonian physics) in order that they should carry weight.

References

1. F.R. Tangherlini, An Introduction to the General Theory of Relativity, Chapter 1: "Special Relativity with a Non-Diagonal Line Element," Suppl. Nuovo Cimento 20, 4-12 (1961).

2. R. Mansouri and R.V. Sexl, "A Test Theory of Special Relativity: I. Simultaneity and Clock Synchronization," General Relativity and Gravitation 8, 497-513 (1977).

3. S. Marinov, Eppur Si Muove (Pierre Libert, Bruxelles, 1977).

4. P.G. Bergmann, Introduction to the Theory of Relativity (Prentice-Hall, Englewood Cliffs, N.J., 1960), pp. 155-156.

5. A. Einstein, Relativity (translation by Robert W. Lawson) (Crown Publishers, New York, 1961), pp. 97-99.

6. A.S. Eddington, The Mathematical Theory of Relativity (Chelsea Publishing Co., New York (3rd ed.), 1975), p. 80; 1st ed. published in 1923.

7. J.L. Anderson, Principles of Relativity Physics (Academic Press, New York, 1967), pp. 181-183.

8. H. Yilmaz, Ann. Phys. 101, 413-432 (1976).

5. COMPUTER SIMULATIONS OF MYELINATED-FIBER REFRACTORY PERIODS

Bell Laboratories (Grant)

Michael H. Brill, Stephen G. Waxman

We have used computer simulations to study relative refractory periods (RRP's) in myelinated fibers, for fibers with uniform and nonuniform nodal and internodal properties. The study examined the relationship between latency variations during the relative refractory period, and conduction path length. The methods used were adapted from those we used previously,[1, 2] and involve integration of differential equations[3, 4] which yield voltage vs time curves at nodes of Ranvier. Stimuli consisted of two 250-μsec current pulses at twice the control threshold. Conduction latencies and average velocities were calculated from the first 50-mV crossings of action potentials at designated points along our simulated 60-node fiber.

The first part of our study concerned RRP's in fibers with identical nodes and internodal regions (uniform fibers). RRP is conventionally defined as the time of recovery of threshold after the absolute refractory period (ARP). Following the suggestion of Raymond and Lettvin[5] that conduction velocity after the ARP has the same time course as threshold recovery, we adopted the convention of defining the RRP as the time during which the conduction velocity of the second spike is less (by a criterion amount) than that of the first spike.

We found that the amplitude of the spike occasioned by the second stimulus is at first significantly less than that occasioned by the first stimulus, but recovers substantially (not completely) by the time the second spike has passed the tenth node of Ranvier.

Also, the latency of the second impulse conducting during the RRP depends on both the interstimulus interval and the propagation distance, being maximal at interstimulus intervals corresponding to the beginning of the RRP, and returning monotonically toward control latency with increasing interstimulus intervals. For any given interstimulus interval, the latency variations are greater for longer conduction distances. The conduction velocity is most slowed at the beginning of the fiber, but increases with increasing conduction distance toward control conduction velocity (that of the first spike, which is 19.7 m/sec). This statement does not imply that the interspike interval also approaches an asymptotic value with increasing distance. We were unable to determine whether the latter asymptote, in fact, exists. Whereas we used a modified Hodgkin-Huxley model of active nerve fiber, models incorporating a period of enhanced conduction speed (a supernormal period) following the RRP cannot have a velocity asymptote without a concomitant latency asymptote. This points up a possible significance of the supernormal period for spike entrainment in long fibers.

In addition to the above study for myelinated fibers with identical nodes and identical internodes, we also performed a refractory-period simulation on inhomogeneous fibers. In these fibers, the first ten nodes were the same as those of the above (fiducial) fiber. The remainder of the fiber (nodes 11-20) was examined for three alternative cases:

a. The nodes had 10 times the area of the fiducial-fiber nodes.

b. The internodes were 75 percent demyelinated compared to those of the fiducial fiber (i.e., myelin thickness was reduced 75 percent, holding constant the axonal diameter).

c. The internode lengths were 2.5 times as long as those of the fiducial fiber.

Uniform fibers with these properties all displayed an increased ARP compared with that of the uniform fiducial fiber. In all three types of nonuniform fiber, we noticed the following effect:

If the second impulse is started near the beginning of what would be the RRP of the fiducial fiber (that is, about 2.1 msec after the first stimulus), the second impulse propagates from one end of the fiber to the other. However, if the second stimulus occurs somewhat later (or is twice the magnitude of the first stimulus), the impulse is blocked when it reaches the inhomogeneity.

This effect, an apparent paradox, can be explained as follows: When the second stimulus is near the beginning of the RRP, there is a utilization time of approximately 1 msec during which the spike develops at the strongly refractory first node before it propagates farther. At subsequent nodes, there is further latency increase relative to control. This latency is large enough for the second part of the fiber to complete its ARP in the wake of the first spike; however, when the interstimulus interval is somewhat longer (or the second stimulus magnitude is larger), the utilization time is shorter,

leading to arrival of the second spike at the inhomogeneity before the second part of the fiber is ready to fire. Thus our simulations provide an example of a situation in which axons may function as high-pass filters. We are now planning to examine the possibility of this effect arising from a cold block in the fiber, in order to assess its further implications.

References

1. M.H. Brill, S.G. Waxman, J.W. Moore, and R.W. Joyner, "Conduction Velocity and Spike Configuration in Myelinated Fibers: Computed Dependence on Internode Distance," J. Neurol., Neurosurg Psychiatry 40, 769-774 (1977).
2. S.G. Waxman and M.H. Brill, "Conduction through Demyelinated Plaques in Multiple Sclerosis: Computer Simulations of Facilitation by Short Internodes," J. Neurol., Neurosurg. Psychiatry 41, 408-416 (1978).
3. A.L. Hodgkin and A.F. Huxley, "A Quantitative Description of Membrane Current and Its Application to Conduction and Excitation in Nerve," J. Physiol. (London) 117, 500-544 (1952).
4. R. Fitzhugh, "Computation of Impulse Initiation and Saltatory Conduction in a Myelinated Nerve Fiber," Biophys. J. 2, 11-21 (1962).
5. S.A. Raymond and J.Y. Lettvin, "Aftereffects in Peripheral Axons as a Clue to Nervous Coding," in S.G. Waxman (Ed.), Physiology and Pathobiology of Axons (Raven Press, New York, 1978), pp. 203-225.

6. FUNCTIONS IN R^m THAT INVERT THEMSELVES AFTER n APPLICATIONS

Bell Laboratories (Grant)

Michael H. Brill

Following the suggestion of D. DeVitt (Perception Technology Corporation; personal communication) that such efforts might be useful in encryption, we examined properties of self-inverting functions – vector-valued functions from R^m to R^m that, when composed n times, give the identity. Some examples for m = 1, n = 2 (functions in one dimension, each its own inverse) are $\pm x^{\pm 1}$, $x/(x-1)$, $\ln(a-e^x)$ ($x < \ln(a)$ in the last case). Examples for higher m and n are n-fold rotations in m-space.

Many of the properties of the case m = 1 generalize to higher m, so we look at the case m = 1 first.

If an invertible function f is its own inverse (n=2), then y = f(x) can be written G(x, y) = 0, where G is symmetric in x and y. [Proof: An expression for the inverse of y = f(x) is x = f(y), in which the variables x and y are reversed. If f is its own inverse, then both of these equations must be true at once, whence $f(x) + f(y) - x - y \equiv G(x, y) = 0$. Here, G is manifestly symmetric in x and y.] An easy way to generate a function f is to find a symmetric differentiable G such that $\partial G/\partial y \neq 0$ and then try

to solve for y.

Now, we show that, if an invertible rational function f is its own inverse (n=2) – except at zeros of its denominator – then f is bilinear. [Proof: Consider a rational function

$$f(x) = \frac{ax^2 + bx + c}{dx^2 + ex + h} = y.$$

This expression cannot be symmetric in x and y unless a = d = 0 (f is bilinear). The result also applies when higher order terms (in x^k) are added to the numerator and denominator. Thus, from the above paragraph, all rational self-inverting functions with m = 1, n = 2 are bilinear.]

What bilinear functions are self-inverting? In one dimension, the general bilinear function is

$$f(x) = \frac{ax + b}{cx + d}.$$

If we define $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ as the characteristic matrix of the transformation (unique up to a scaling factor p), we can show that composing bilinear functions corresponds to multiplying their characteristic matrices. Clearly, then, a bilinear function is self-inverting after n applications if we can choose p such that the eigenvalues of its characteristic matrix are distinct n^{th} roots of unity. (When the matrix is nondiagonal, the eigenvalues must be distinct. This can be seen from the counterexample y = x + c, which is not self-inverting even though its characteristic matrix has two eigenvalues +1.)

What other functions are self-inverting? Let f be a bilinear function that inverts itself after n = 2 applications. Let g be an invertible function whose domain and range span the real line. Then $g^{-1} \circ f \circ g$ will be self-inverting. If we are careful to restrict the domain x (which must anyway be restricted so the denominator of f is nonzero), we can enlarge the class of g by removing the restriction on the range. Thus, in the above example $\ln(a-e^x)$, x must be less than ln (a). For n > 2, the domains of f and g must, of course, be further restricted so $(g^{-1} \circ f \circ g)^k$ is invertible for $k \leq n$. We conjecture (but have not proved) that the above constructive method will generate all self-inverting functions for m = 1.

The findings of the last two paragraphs can be generalized to all m > 0. The bilinear transformation becomes a projective transformation (each of whose components has the same denominator), and the characteristic matrix (a_{ij}) has dimension (m+1) × (m+1). The bilinear function $\underline{f}$ has m components

$$f_i(\underline{x}) = \frac{a_{i1}x_1 + \ldots + a_{im}x_m + a_{i,m+1}}{a_{m+1,1}x_1 + \ldots + a_{m+1,m}x_m + a_{m+1,m+1}} \qquad (i = 1, \ldots, m).$$

The eigenvalue condition carries over from the case m = 1; however, the eigenvalues need not be distinct so long as the matrix (a_{ij}) is diagonalizable (i.e., it is in block-diagonal form such that the eigenvalues in each block are distinct).

A straightforward generalization of g defines vector-valued $\underline{g}$.

An example for m = 2 is $(x^a y^b, x^c y^d)$, where the characteristic matrix is chosen as $\begin{pmatrix} a & b & 0 \\ c & d & 0 \\ 0 & 0 & 1 \end{pmatrix}$, $\underline{g}$ is the logarithm (in any base) applied to the two coordinates, and one must insure $x, y > 0$. When n = 2, the eigenvalue conditions correspond to $a + d = 0$, $a^2 + bc = 1$. Two of the eigenvalues are the same here, but not in the same matrix block.

Since any projective transformation $\underline{f}$ is expressible as $\underline{F}^{-1} \circ \underline{L} \circ \underline{F}$ (where $\underline{F}$ is a projective transformation and $\underline{L}$ is an affine transformation), the above theory can be recast about the affine transformation. $\underline{L}(\underline{x})$ has components of the form $L_i(\underline{x}) = a_{i1}x_1 + \ldots + a_{im}x_m + a_{i,m+1}$. The characteristic matrix is defined as before, but the bottom row is $(0, \ldots, 0, 1)$, and there are no singularities in $\underline{L}$.

A note about projective transformations as they might apply to vision:

Cornsweet[1] has noted that, as a function of light intensity x, the visual response (at the receptor level) is $\frac{ax}{b + x}$. Since this is a one-dimensional projective transformation, it supports the 4-point invariant

$$\left| \frac{(x_1 - x_2)(x_3 - x_4)}{(x_1 - x_3)(x_2 - x_4)} \right|.$$

This may be useful in understanding how the visual system can deduce brightness relationships independently of the constants (adaptation state) of the transduction function.

References

1. T. Cornsweet, Visual Perception (Academic Press, New York, 1970).

7. NEW TEST PATTERNS FOR CAMERA LENS EVALUATION

National Institutes of Health (Training Grant 5 TO1 EY00090)

Bradford Howland

The method preferred today for testing camera lenses is the direct measurement of the modulation transfer function (mtf) using specialized electronic instruments.[1] Photographic methods for making similar measurements are less accurate, and usually require the use of a microdensitometer. An attempt to circumvent this limitation is

that of Putora, who photographs circular test patterns of varying fineness with high-contrast film; the lens resolution is determined directly by inspection of the negative.[1] We describe here additional test patterns which provide direct indication of lens performance when photographed with high-contrast film.

Figure XXV-1 shows two of the new charts, as follows: (a) the vernier resolution chart, and (b) the triangular-wave pattern. The vernier chart utilizes pairs of black lines, displaced by successively increasing increments, on a 20% grey background. White borders to the black lines are placed as indicated. The action of the chart depends on the small differential motions of the images of the black lines due to the spreading of the white into the black, and the black into the grey when an unsharp image is rendered. The degree of unsharpness of the image is indicated by the altered vernier correspondence of the displaced line images.

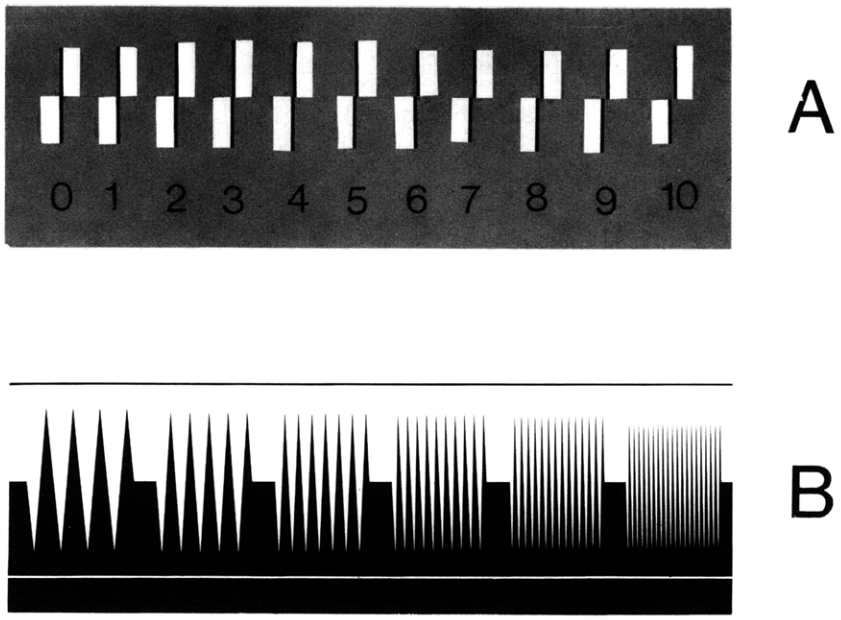

Fig. XXV-1. Two new charts for the measurement of lens resolution: (A) the vernier resolution chart, (B) the triangular-wave pattern.

This behavior is illustrated in Fig. XXV-2, which shows photographs of the chart with successive stages of defocus – note that the vernier correspondence, which occurs at number 0 in the original (A), occurs at line pairs numbered 3, 8, and 10 in the progressively defocused images B, C, and D. Here we are using a defocused image as a simulation of the lens defects of astigmatism, or curvature of field. Tests indicate that defocus, lens aberrations, and diffraction blur all affect the vernier sharpness index.

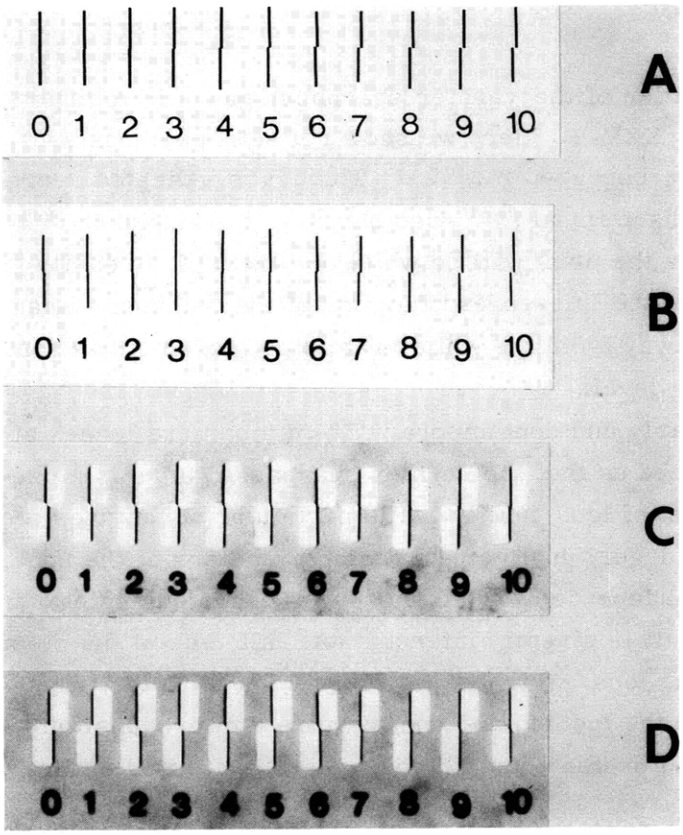

Fig. XXV-2.

High-contrast photographs showing (A) sharp imagery, and (B, C, and D) successive stages of defocus of the vernier resolution chart. Note that vernier correspondence occurs at line pairs marked 0 in frame A, 3 in frame B, 7 in frame C, and 10 in frame D.

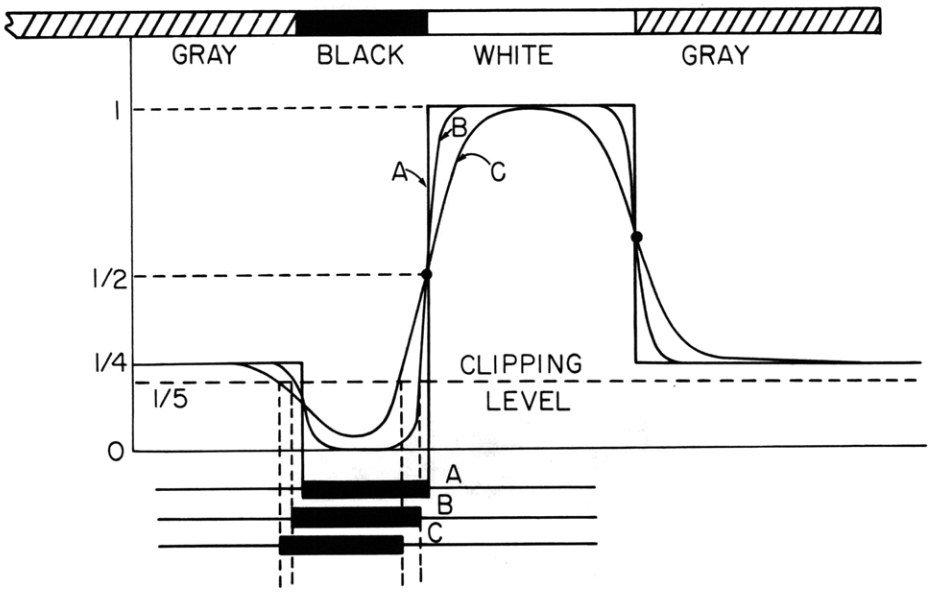

Fig. XXV-3. Reflectance cross sections of (A) sharp, and (B, C) unsharp images of lines forming the vernier resolution chart.

A qualitative explanation for the action of the vernier sharpness gauge is facilitated by consideration of the diagram of Fig. XXV-3. Here we show reflectance cross sections of images of the lines of the chart, bordered in white. The cases indicated represent perfect, slightly unsharp, and unsharp imagery. Also shown is the assumed 20% clipping level which determines whether the print will be white or black. It is easily seen that the result of the finite slope of the line-spread function of the defocused images is to move the image of the black line away from the white area; this increment is somewhat dependent on the choice of clipping level.

Extensive tests with a variety of early and contemporary 35-mm camera lenses at differing aperture settings have convinced us that the vernier sharpness gauge is a useful and sensitive test for the assessment of lens quality. The particular advantages of this method are: (a) the test result is largely unaffected by the parameters of the photographic process, (b) the test result is independent of the sharpness of the enlarging lens used to make the print, (c) the test result is directly interpretable in terms of the steepness of the edge response function of the lens, which is, of course, the Fourier transform of the mtf response, and (e) since the method does not require a microdensitometer, it can be used both by amateur and professional photographers of limited budget or little scientific training.

A second test method is the use of the triangular-wave pattern shown in Fig. XXV-1(b). For this purpose, a set of patterns of geometrically increasing spatial frequency was generated on an oscilloscope and combined as a photo-montage. In Fig. XXV-4(a) and 4(b) we show the result of photographing this montage with an excellent and an inferior lens, with high-contrast film, using an exposure index chosen to render the pattern symmetrically about the white-black axis. We note that the envelope of the shaded areas renders an approximate plot of the mtf response of the lens; if the response is highly degraded, this approximation becomes quite exact. Analysis shows that, for only slightly degraded responses, the falloff in the height of the triangular-wave pattern is proportional to the linear extent of the edge response function, evaluated

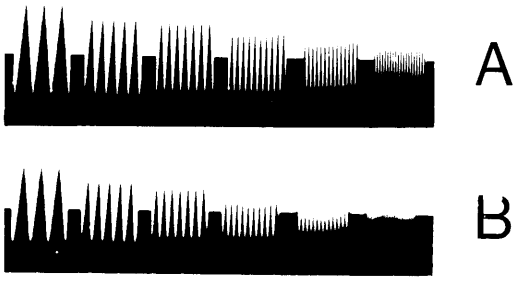

Fig. XXV-4. High-contrast renditions of the triangular-wave patterns photographed at a distance using (A) a high-quality lens, and (B) an inferior lens.

at the 25% and 75% points.

These results suggested the possible existence of a pattern which, when photographed with contrast ratio approaching infinity, would indicate the exact mtf

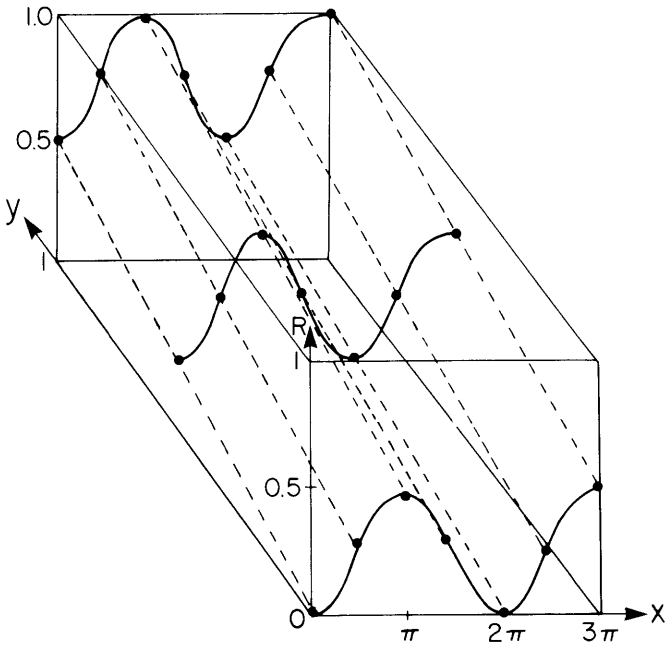

Fig. XXV-5. Two-dimensional pattern of reflectance for a new form of mtf measurement chart. High-contrast photographs of an ensemble of such patterns of increasing spatial frequency render an exact plot of the mtf function.

response of the lens. In Fig. XXV-5 we show the realization of such a pattern, which requires continuous tone reflectance variation according to the equation:

$$R = \frac{1}{4} \sin \omega x + \frac{1}{2} y + \frac{1}{4} \qquad 0 \leq R \leq 1, \quad 0 \leq y \leq 1.$$

In this figure, the scale of the x-coordinate has been enlarged for clarity. The reflectance exhibits a sinusoidal variation with x, together with a linear shading with y. The operation of this chart is explained by analogy with the sphygmomanometer, used to measure blood pressure. The reflectance at the top of the chart varies from 50% to 100%, at the middle from 25% to 75%, and at the bottom from 0% to 50%. Thus, if the film clips at 50% reflectance, all levels of the chart will contain both black and white areas, and the sinusoid will be reproduced in silhouette. Now, however, if the sinusoidal component of reflectance is reduced during the imagery by a factor, $\alpha < 1$, then the regions

below the line $y = \alpha/2$ will be solid white, and above the line $y = 1 - \alpha/2$ will be black. The extent of the intermediate, shaded region will therefore be reduced by the same factor α. A composite chart, reminiscent of Fergus Campbell's variable-contrast mtf chart, featuring progressively increasing spatial frequencies will, when so photographed, render an exact plot of the mtf response of the lens. Electronic circuits to facilitate the generation of this new test pattern are under development.

A possible further application of the vernier chart is to measure the performance of microscope lenses. The advantage here is that the vernier chart does not require line elements too fine to reproduce by microphotography, as is the case with conventional resolution charts. One further application of these methods is measurement of the resolution of cathode ray tubes. We have designed and constructed TTL-logic circuitry to generate, at varying clock rates, the triangular-wave patterns, the vernier resolution chart, and an additional pattern useful for the detection of coma. These patterns easily demonstrate the limitations of available cathode ray tubes.

References

1. L.R. Baker and T. Moss,"Electro-Optical Methods of Image Evaluation," Electro-Optical Systems Design Conference, New York, N.Y., September 1969.

2. I. Putora, "The Sharpness Indicator," J. Soc. Motion Picture Television Engrs. 78, 956-960 (1969).