Proc. Natl. Acad. Sci. USA
Vol. 87, pp. 7819–7823, October 1990
Biophysics

Structure of visual perception

(motion detector/perceptual "oneness"/non-Euclidean space/geodesic/Hering illusion)

JUN ZHANG*† AND SIYE WU‡§

*Neurobiology Group, University of California, Berkeley, CA 94720; and ‡Department of Mathematics, Massachusetts Institute of Technology, Cambridge, MA 02139

Communicated by Russell L. De Valois, July 9, 1990 (received for review December 19, 1989)

ABSTRACT **The response properties of a class of motion detectors (Reichardt detectors) are investigated extensively here. Since the outputs of the detectors, responding to an image undergoing two-dimensional rigid translation, are dependent on both the image velocity and the image intensity distribution, they are nonuniform across the entire image, even though the object is moving rigidly as a whole. To achieve perceptual "oneness" in the rigid motion, we are led to contend that visual perception must take place in a space that is non-Euclidean in nature. We then derive the affine connection and the metric of this perceptual space. The Riemann curvature tensor is identically zero, which means that the perceptual space is intrinsically flat. A geodesic in this space is composed of points of constant image intensity gradient along a certain direction. The deviation of geodesics (which are perceptually "straight") from physically straight lines may offer an explanation to the perceptual distortion of angular relationships such as the Hering illusion.**

To understand the perceptual organization of the primate visual system has always been a challenging endeavor for the inquiring human mind. Major advances have been made since, among others, the pioneering work of Hubel and Wiesel (1–3), which established that neurons in the primary visual cortex (V1) of mammals respond selectively to bars and edges of a restricted range of orientations. Recently, convincing evidence began to accumulate, both from physiological/anatomical probes and clinical/behavioral studies, that supports the long-proclaimed hypothesis of segregation of different subsystems in visual processing at various stages, notably the parvocellular system, which might be related to form processing, and the magnocellular system, which may be involved with motion processing (for review, see refs. 4–8). Despite this functional subdividing, there are constant interactions between the two systems, contributing to the final percept of either form or motion. For example, one can acquire the sense of motion simply by presenting a visual stimulus at two different locations in succession with appropriate time delay, the phenomenon of apparent motion. On the other hand, one can identify vividly the form of a textural figure moving against a background of the same texture that is otherwise unidentifiable, a clear indication that form perception can be derived from pure motion information. In this paper, the intrinsic structure of the motion system will be extensively investigated and its intrusion to the form perception, briefly discussed from a theoretical point of view. We will show that motion perception, or visual perception in general, can be viewed as an interpretation of sensory data based on an intrinsic geometry that in turn is determined by rules of organizing the sensory data.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Motion Detectors and Their Properties

It is generally believed that the motion system incorporates some structural elements, called motion detectors, to preprocess an image. A prototypical detector, proposed by Reichardt, Hassenstein, and their colleagues in the insect visual system (9, 10), performs two-point correlations of a visual image and was later proved to be the building block for the general scheme of motion computation with n-point inputs (11). It has since been shown that when the filters of these so-called Reichardt detectors are elaborated with spatial–temporal bandpass properties, they may extract the Fourier power spectrum of the visual image based on their filtering characteristics (12, 13). In this way, they resemble motion energy detectors, which have more or less been correlated to both the neurophysiological properties of cortical neurons and human psychophysical performance (14–18). On the other hand, it has been shown only recently that, if the spatial separation and the temporal delay of the filters can be regarded as infinitesimally small, Reichardt detectors may actually extract the local velocity vectors based on the spatial and temporal gradient of an image (19). It was demonstrated that, for the simplest case of an image undergoing two-dimensional rigid, translatory motion with velocity $\mathbf{v} = [v_1(t), v_2(t)]^T$ (here and throughout the paper $[\cdot, \cdot]^T$ denotes vector transpose), the response of Reichardt detectors $\mathbf{V} = [V_1(x, y, t), V_2(x, y, t)]^T$ is

$$\begin{bmatrix} V_1 \\ V_2 \end{bmatrix} = \epsilon e^{-2q} \begin{bmatrix} q_{xx} & q_{xy} \\ q_{yx} & q_{yy} \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}, \qquad [1]$$

where $f(x(t), y(t))$ is the image intensity function, ϵ is a constant, $q(x, y) = \log f(x, y)$, and $q_{xx} = \partial^2 q/\partial x^2$, etc. This form resembles the response of motion field detectors that perform spatial (∇) and temporal ($\partial/\partial t$) differentiations in successive stages (20, 21)

$$\mathbf{V} = \frac{\partial}{\partial t} \nabla f, \qquad [2]$$

or, when explicitly written out,

$$\begin{bmatrix} V_1 \\ V_2 \end{bmatrix} = \begin{bmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}. \qquad [3]$$

Discarding the overall scaling of the detector response and the nonlinear (logarithmic) compression of the image intensity in Eq. **1** which, by intuition, are less important, we shall explore the consequences of the vectorial relationship as expressed by Eq. **3** in the rest of the paper.

†To whom reprint requests should be addressed at: Department of Psychology, 3210 Tolman Hall, University of California, Berkeley, CA 94720.
§Present address: Department of Mathematics, Columbia University, New York, NY 10027.

We first convert Eq. **3** into a tensorial equation [i.e., an equation that is invariant in form under coordinate transformations $\bar{x}^\lambda = \bar{x}^\lambda(x^\sigma)$] by restricting ourselves to the general linear transformation (with constant Jacobian)

$$\frac{\partial \bar{x}^\lambda}{\partial x^\sigma} = \text{constants.} \qquad [4]$$

In this case, Eq. **3** can be compactly expressed as¶

$$V_\alpha = f_{\alpha\beta} v^\beta, \qquad [5]$$

where $f_{\alpha\beta}$ denotes the second-order spatial derivatives, known as the Hessian of $f(x, y)$, and α, β run over 1, 2 (i.e., x, y). With the condition of constant Jacobian, $f_{\alpha\beta}$ is indeed a tensor of rank 2 under coordinate transformations; V_α and v^β are covariant and contravariant vectors, respectively.

The response of motion detectors V_α given by Eq. **3** or Eq. **5** has the following desirable properties: (*i*) It is covariant; i.e., the response of any one of the detectors in a family at a specific location represents the same vector regardless of the choice of a coordinate axis. This is of fundamental importance for a biological system. Each motion detector in the visual system is made up of a pair of receptors (subunits) with appropriate separation in space and delay in time (both parameters are regarded as infinitesimally small and related to the constant ϵ in Eq. **1**). Because of the random nature of the neuronal distribution, the direction of any one of the many pairs of receptors and therefore the axis ("preferred" direction) of each detector is essentially random, covering a range of 360°. It is crucial that all detectors at a given location give mutually compatible signals about the local image velocity v^β, regardless of the individual "label" of their preferred directions. This is possible only when V_α is covariant under a rotational coordinate transformation (a special case of the general linear transformations of Eq. **4**). (*ii*) It is complete; i.e., the local image velocity v^β can be unambiguously determined by the detector response V_α (by inverting the matrix $f_{\alpha\beta}$). The only exception occurs when the determinant of $f_{\alpha\beta}$ is zero or $f_{xx}f_{yy} - f_{xy}^2 = 0$. The corresponding image intensity profiles would, either locally or at large, assume the shape of developable surfaces. The situation here is closely related to the so-called "aperture problem" (22–24)—namely, it is impossible to specify the image velocity along the direction corresponding to the ruling of the surface. This is a restriction set by the physical nature of an image instead of the biological processing of it. Any physically extractable information about image motion is faithfully preserved in the response of the detectors.

Fundamental Problem of Rigid Motion

The response of motion detectors V_α confounds image velocity information with image intensity information. As given by Eq. **5**, V_α is dependent on the image velocity v^β as well as the Hessian $f_{\alpha\beta}$ of the image intensity function for any two-dimensional translatory motion. Even though an object moves rigidly (i.e., v^β is not a function of x and y), V_α is still a function of the spatial location x, y, since $f_{\alpha\beta}$ are generally not constant over the space. In other words, the response of motion detectors is not uniform even though each and every point of the object moves with uniform velocity (rigid motion). Therefore one question naturally arises: how then could we (as human beings) ever derive the percept that the object is moving rigidly as a whole entity and in its entirety, as opposed to some nonrigid, irregular deformations? How could we perceive the "oneness" in a moving object? To pose this question in mathematical language, we would like to know how a nonuniform vector field V_α could ever be construed as a constant vector field.

Non-Euclidean Perceptual Space

At first glance, it may appear that it would never be possible to identify a nonuniform vector field with a constant vector field. However, if we introduce the notion of non-Euclidean geometry, there is a natural solution to this problem. Recall that in a general affine space, the constancy of a vector field at different spatial locations is not defined by simply comparing each of their components. Instead, it is achieved by "transplanting" the vector over the space and at the same time changing the vector components according to rules carefully designed so that the generalized notion of vector parallelism is still preserved during this transplantation. By transporting a vector parallel from one location to another, the comparison of vectors remains meaningful and faithful. This law of parallel transplantation involves a set of coefficients denoted by $\Gamma^\lambda_{\alpha\beta}$, the so-called coefficients of affine connection or simply connection, which unequivocally characterizes the affine structure of the space (an affine space is a space that carries invariant structure under a linear transformation of coordinates). Therefore in an affine space, two vectors at different spatial locations are considered parallel (or "equal") if one can be transported to the location of and then coincide with the other following the prescribed transplantation law. This piece of knowledge led us to propose that the perceptual space, a space that is a mapping from the physical space and where "perception" takes place, is nonetheless a generalized affine space with its particular affine connection. From this viewpoint, V_α *is* constant over the perceptual space with some $\Gamma^\lambda_{\alpha\beta}$ (yet to be specified), though it is *not* uniform in the physical space, as we pointed out earlier. In fact, the constancy of V_α in the perceptual space can be used in this way to define the coefficients of affine connection $\Gamma^\lambda_{\alpha\beta}$.‖

Affine Connection. Recall that in the affine space, an arbitrary vector ξ_α is transported along any curve $x^\beta(s)$ from a point s to a neighboring point $s + ds$ according to the following law of vector transplantation:

$$\frac{d\xi_\alpha}{ds} = \Gamma^\lambda_{\alpha\beta}\xi_\lambda \frac{dx^\beta}{ds}. \qquad [6]$$

Eq. **6** defines a "constant" vector field along the curve $x^\beta(s)$. Note that it is bilinear in ξ_α and dx^β. In the present case, we require the response of all motion detectors V_α to be constant in the space over the region that the object occupies. The change of V_α along the curve $x^\beta(s)$ is

$$\frac{dV_\alpha}{ds} = \frac{\partial V_\alpha}{\partial x^\beta}\frac{dx^\beta}{ds} = \left(\frac{\partial f_{\alpha\rho}}{\partial x^\beta}\right) v^\rho \frac{dx^\beta}{ds}, \qquad [7]$$

where we have used Eq. **5** and the fact that $\partial v^\rho/\partial x^\beta = 0$ (since v^ρ is constant in the physical space). We temporarily assume that, at the points in consideration, the matrix $f_{\alpha\rho}$ is nonsingular—namely, its determinant $\mathcal{F}$ is not zero (we call the points with $\mathcal{F} = 0$ *degenerating points* in this paper). Define the inverse of $f_{\alpha\rho}$ as $f^{\rho\alpha}$ using the Kronecker δ^β_α; i.e.,

$$f_{\alpha\rho} f^{\rho\beta} = \delta^\beta_\alpha. \qquad [8]$$

¶Throughout this paper, the Einstein summation convention is respected unless otherwise noted. This convention is as follows: if an index appears twice in an expression, once as an upper index and once as a lower index, then a summation is implied over that index.

‖The mathematical background of this section can be found in standard textbooks on differential geometry or general relativity (see, for example, ref. 25).

Then we may solve for v^ρ from Eq. 5 and express it as

$$v^\rho = f^{\rho\lambda} V_\lambda. \quad [9]$$

Inserting Eq. 9 into Eq. 7, we have

$$\frac{dV_\alpha}{ds} = \frac{\partial f_{\alpha\rho}}{\partial x^\beta} f^{\rho\lambda} V_\lambda \frac{dx^\beta}{ds}. \quad [10]$$

Comparing the forms of Eq. 6 and Eq. 10, we immediately derive the affine connection of the perceptual space

$$\Gamma^\lambda_{\alpha\beta} = \frac{\partial f_{\alpha\rho}}{\partial x^\beta} f^{\rho\lambda} \quad [11]$$

or, equivalently,

$$f_{\rho\lambda}\Gamma^\lambda_{\alpha\beta} = \frac{\partial f_{\alpha\rho}}{\partial x^\beta}. \quad [12]$$

Eq. 11 or 12 defines the connection of the non-Euclidean perceptual space, and thus is of fundamental importance. $\Gamma^\lambda_{\alpha\beta}$ is symmetric with respect to its lower indices: $\Gamma^\lambda_{\alpha\beta} = \Gamma^\lambda_{\beta\alpha}$. Since any affine space should be characterized by its connection $\Gamma^\lambda_{\alpha\beta}$, we expect that the structure of visual perception would be revealed as a consequence of our assertion of the perceptual space as being a general affine space. In particular, since the visual space is a metric space, we expect that the metrical properties would be derived from Eq. 11.

Riemann Curvature Tensor. Before discussing the metrical properties of the perceptual space, let us first derive its Riemann curvature tensor $R^\alpha_{\eta\beta\gamma}$. By definition,

$$R^\alpha_{\eta\beta\gamma} = \frac{\partial \Gamma^\alpha_{\beta\eta}}{\partial x^\gamma} - \frac{\partial \Gamma^\alpha_{\eta\gamma}}{\partial x^\beta} + \Gamma^\alpha_{\tau\gamma}\Gamma^\tau_{\beta\eta} - \Gamma^\alpha_{\tau\beta}\Gamma^\tau_{\gamma\eta}. \quad [13]$$

PROPOSITION 1. *For the perceptual space whose connection is given by Eq.* 11, *the Riemann curvature tensor is*

$$R^\alpha_{\eta\beta\gamma} = 0. \quad [14]$$

Proof. First notice the fact that $f_{\alpha\beta}$ is nothing but $\partial^2 f/\partial x^\alpha \partial x^\beta$. This means that we can regard *lower* indices of f both as tensor indices and as differentiation indices and thus interchange their orders at will, simply due to the commutativity of differentiations with respect to coordinates x and y. Substituting Eq. 11 into Eq. 13, differentiating, and collecting terms, we have

$$R^\alpha_{\eta\beta\gamma} = (f^{\alpha\rho}_{,\gamma} + f^{\alpha\mu}f^{\tau\rho}f_{\tau\mu\gamma})f_{\beta\rho\eta} - (f^{\alpha\rho}_{,\beta} + f^{\alpha\mu}f^{\tau\rho}f_{\tau\mu\beta})f_{\gamma\rho\eta}, \quad [15]$$

where we use $,\gamma$ to denote coordinate differentiations such that $f^{\alpha\rho}_{,\gamma} = \partial f^{\alpha\rho}/\partial x^\gamma$, etc., and of course $f_{\tau\mu,\gamma} = f_{\tau\mu\gamma}$, etc. Using the following identity obtained by differentiating Eq. 8:

$$f^{\alpha\rho}_{,\gamma}f_{\rho\beta} = -f^{\alpha\rho}f_{\rho\beta,\gamma}, \quad [16]$$

the terms within the first parentheses of Eq. 15 become

$$f^{\alpha\rho}_{,\gamma} + f^{\alpha\mu}f^{\tau\rho}f_{\tau\mu\gamma} = f^{\alpha\rho}_{,\gamma} + f^{\alpha\mu}(-f_{\tau\mu}f^{\tau\rho}_{,\gamma}) = f^{\alpha\rho}_{,\gamma} - \delta^\alpha_\tau f^{\tau\rho}_{,\gamma} = 0. \quad [17]$$

Similarly the terms within the second parentheses yield zero as well. This proves that the Riemann tensor $R^\alpha_{\eta\beta\gamma}$ is identically zero. q.e.d.

The fact that the Riemann curvature tensor is everywhere zero (except for those degenerating points where $f^{\alpha\rho}$ does not exist) means that the perceptual space is, after all, *flat* or pseudo-Euclidean—there is no intrinsic curveness. Since all components of $R^\alpha_{\eta\beta\gamma}$ are zero (although in the two-dimensional case, there is only one independent component), the Ricci tensor $R_{\eta\gamma}$ (given by contracting α, β in $R^\alpha_{\eta\beta\gamma}$) is identically zero, and hence the Gaussian curvature is also zero.

Proposition 1 has the following important consequence that is perceptually justifiable. Normally, the transplantation of a vector between any two end-points, according to the law of parallel transplantation (Eq. 6), is dependent on the path connecting the two points. Indeed, if an arbitrary vector ξ_η at one point is parallel transported along segments of two curves, dx^β and $d\tilde{x}^\gamma$, sequentially but in different order, to the same end-point, then the difference of changes in the vector component $\Delta\xi_\eta$ following these two paths of transportation is given by

$$\Delta\xi_\eta = -R^\alpha_{\eta\beta\gamma}\xi_\alpha dx^\beta d\tilde{x}^\gamma. \quad [18]$$

Since in the present case $\Delta\xi_\eta = 0$, this means that the parallel transportation of vectors between any two points is path independent and hence uniquely defined. In terms of motion perception, this implies that the comparison of local velocities among all points in the space can be performed unambiguously even at large, that is, with finite separations.

Geodesic. We may study the geodesics of an affine space without referring to its metrical properties, if we deliberately choose not to. In this case, the geodesic equation defines a curve that has zero geodesic curvature (the generalized notion of a "straight" line in an affine space)

$$\frac{d^2x^\lambda}{ds^2} + \Gamma^\lambda_{\alpha\beta}\frac{dx^\alpha}{ds}\frac{dx^\beta}{ds} = 0, \quad [19]$$

where s is some parameter (not necessarily arc length since we have not defined length yet).

PROPOSITION 2. *The geodesics of the perceptual space can be explicitly expressed as*

$$f_\sigma = b_\sigma s + c_\sigma, \quad [20]$$

where $f_\sigma = \partial f/\partial x^\sigma$ *and the two sets of constants* b_σ, c_σ *determine the direction and the initial position of a particular geodesic.*

Proof. By the definition of $\Gamma^\lambda_{\alpha\beta}$ in Eq. 11, the geodesic equation (Eq. 19) becomes

$$\frac{d^2x^\lambda}{ds^2} + \frac{\partial f_{\alpha\rho}}{\partial x^\beta} f^{\rho\lambda}\frac{dx^\alpha}{ds}\frac{dx^\beta}{ds} = 0. \quad [21]$$

Multiplying by $f_{\sigma\lambda}$, we have

$$f_{\sigma\lambda}\frac{d^2x^\lambda}{ds^2} + \frac{\partial f_{\alpha\sigma}}{\partial x^\beta}\frac{dx^\beta}{ds}\frac{dx^\alpha}{ds} = 0. \quad [22]$$

The left-hand side turns out to be an exact differentiation

$$\frac{d}{ds}\left(f_{\sigma\alpha}\frac{dx^\alpha}{ds}\right) = 0, \quad [23]$$

which can be simplified to

$$\frac{d^2 f_\sigma}{ds^2} = 0. \quad [24]$$

Therefore the solution to the geodesic equation is given by Eq. 20. q.e.d.

There always exists a direction (with direction numbers l^σ) orthogonal to the vector b_σ, so that

$$l^\sigma f_\sigma = l^\sigma(b_\sigma s + c_\sigma) = l^\sigma c_\sigma = \text{constant}. \quad [25]$$

Eq. **25** shows that a geodesic is composed of points that have constant image intensity gradient along direction l^σ. Drawn in the two-dimensional visual space, geodesics are almost always different from straight lines of the physical space.

Metric Tensor. Finally, we come to the discussion of the metric tensor $g_{\mu\nu}$ of the perceptual space. Note that $g_{\mu\nu}$ is related to $\Gamma^\lambda_{\alpha\beta}$ by the following identity:

$$\Gamma^\lambda_{\mu\nu} = \frac{g^{\lambda\sigma}}{2}\left(\frac{\partial g_{\mu\sigma}}{\partial x^\nu} + \frac{\partial g_{\nu\sigma}}{\partial x^\mu} - \frac{\partial g_{\mu\nu}}{\partial x^\sigma}\right). \qquad [26]$$

It is common knowledge that our visual space is a metric space, in that we can make estimations and judgments about length and separation. Yet the exact relationship of this visual metric and the metric of the perceptual space is still under investigation; they are certainly not identical. Some preliminary results show that they might be crucial in the coordination of the perceptual space, which is out of the scope of this paper. Nevertheless, it would be interesting to see whether, if Eqs. **11** and **26** are combined to solve for $g_{\mu\nu}$, one can obtain a closed solution in terms of $f_{\alpha\beta}$.

Contracting α and λ in Eq. **11**, we have

$$\Gamma^\lambda_{\lambda\beta} = \frac{\partial f_{\lambda\rho}}{\partial x^\beta} f^{\rho\lambda}. \qquad [27]$$

Using the calculation of the matrix inverse by its cofactors in the determinant $\mathcal{F}$, i.e.,

$$f^{\rho\lambda} = \frac{1}{\mathcal{F}}\frac{\partial \mathcal{F}}{\partial f_{\lambda\rho}}, \qquad [28]$$

we find the right-hand side of Eq. **27** to be

$$\frac{\partial f_{\lambda\rho}}{\partial x^\beta} f^{\rho\lambda} = \frac{\partial f_{\lambda\rho}}{\partial x^\beta}\left(\frac{1}{\mathcal{F}}\frac{\partial \mathcal{F}}{\partial f_{\lambda\rho}}\right) = \frac{1}{\mathcal{F}}\frac{\partial \mathcal{F}}{\partial x^\beta} = \frac{\partial \log|\mathcal{F}|}{\partial x^\beta}. \qquad [29]$$

On the other hand, the left-hand side of Eq. **27** is the well-known result

$$\Gamma^\lambda_{\lambda\beta} = \frac{1}{2}\frac{\partial \log|\mathcal{G}|}{\partial x^\beta}, \qquad [30]$$

where $\mathcal{G}$ is the determinant of the metric tensor $g_{\mu\nu}$. Combining Eqs. **29** and **30**, we have

$$\frac{\partial \log|\mathcal{F}|}{\partial x^\beta} = \frac{1}{2}\frac{\partial \log|\mathcal{G}|}{\partial x^\beta} \qquad [31]$$

or

$$\frac{\partial}{\partial x^\beta}(\log|\mathcal{G}| - \log\mathcal{F}^2) = 0. \qquad [32]$$

Hence

$$\mathcal{G} = (\varepsilon\mathcal{F})^2, \qquad [33]$$

with constant ε relating the measuring unit of the perceptual space to that of the physical space. For convenience, we set $\varepsilon = 1$. To express the matrix element $g_{\mu\nu}$ in terms of $f_{\alpha\beta}$, an educated guess based on Eq. **33** and its proof would be the following.

PROPOSITION 3. *The metric tensor that satisfies Eq.* **11** *is given by*

$$g_{\mu\nu} = \sum_\alpha f_{\mu\alpha}f_{\alpha\nu}, \qquad [34]$$

where the summation over α is enforced.

Proof. Clearly, the inverse of $g_{\mu\nu}$ obeying Eq. **34** would be

$$g^{\lambda\sigma} = \sum_\beta f^{\lambda\beta}f^{\beta\sigma}, \qquad [35]$$

so that the following orthogonality condition is satisfied:

$$g^{\mu\lambda}g_{\lambda\nu} = \left(\sum_\beta f^{\mu\beta}f^{\beta\lambda}\right)\left(\sum_\alpha f_{\lambda\alpha}f_{\alpha\nu}\right) = \sum_{\alpha,\beta} f^{\mu\beta}f_{\alpha\nu}(f^{\beta\lambda}f_{\lambda\alpha})$$
$$= \sum_{\alpha,\beta} f^{\mu\beta}f_{\alpha\nu}\delta^\beta_\alpha = \sum_\alpha f^{\mu\alpha}f_{\alpha\nu} = \delta^\mu_\nu. \qquad [36]$$

From Eq. **26**

$$\Gamma^\lambda_{\mu\nu} = \frac{1}{2}\sum_\beta f^{\lambda\beta}f^{\beta\sigma}\Bigg(\sum_\alpha f_{\mu\alpha,\nu}f_{\alpha\sigma} + \sum_\alpha f_{\mu\alpha}f_{\alpha\sigma,\nu} + \sum_\alpha f_{\nu\alpha,\mu}f_{\alpha\sigma}$$
$$+ \sum_\alpha f_{\nu\alpha}f_{\alpha\sigma,\mu} - \sum_\alpha f_{\mu\alpha,\sigma}f_{\alpha\nu} - \sum_\alpha f_{\mu\alpha}f_{\alpha\nu,\sigma}\Bigg)$$
$$= \frac{1}{2}\sum_{\alpha,\beta} f^{\lambda\beta}f^{\beta\sigma}(f_{\mu\alpha,\nu}f_{\alpha\sigma} + f_{\nu\alpha,\mu}f_{\alpha\sigma})$$
$$= \frac{1}{2}\sum_\beta f^{\lambda\beta}f_{\mu\beta,\nu} + \frac{1}{2}\sum_\beta f^{\lambda\beta}f_{\nu\beta,\mu} = f^{\lambda\beta}f_{\mu\nu\beta}. \qquad [37]$$

This proves that $g_{\mu\nu}$ given by Eq. **34** indeed satisfies Eq. **11**, our original definition of the connection of the perceptual space. q.e.d.

The $g_{\mu\nu}$ thus obtained is symmetric with respect to its indices (note that their subscripts are, unlike the subscripts of $f_{\alpha\beta}$, solely tensor indices). Moreover, since $g_{11} = f^2_{xx} + f^2_{xy} \geq 0$, $\mathcal{G} = \mathcal{F}^2 \geq 0$, $g_{\mu\nu}$ is semipositive definite. Therefore it has all the desired properties of being a metric tensor. The only points where $\mathcal{G} = 0$ is where $\mathcal{F} = 0$ or $f_{xx}f_{yy} - f^2_{xy} = 0$, the degenerating points, as we shall discuss below.

Degenerating Points. At the degenerating points where $\mathcal{F} = 0$, $f^{\alpha\beta}$ is not defined. However, the metric $g_{\mu\nu}$ given by Eq. **34** is still determined. Since $f_{\alpha\beta} = f_{\beta\alpha}$, one can always diagonalize $f_{\alpha\beta}$ and hence the metric tensor $g_{\mu\nu}$. The eigenvalues of $f_{\alpha\beta}$ are given by solving

$$\begin{vmatrix} f_{xx} - \lambda & f_{xy} \\ f_{yx} & f_{yy} - \lambda \end{vmatrix} = 0 \qquad [38]$$

or

$$\lambda^2 - (f_{xx} + f_{yy})\lambda + f_{xx}f_{yy} - f^2_{xy} = 0. \qquad [39]$$

The two eigenvalues are therefore expressed as

$$\lambda_{1,2} = [(f_{xx} + f_{yy}) \pm \sqrt{(f_{xx} - f_{yy})^2 + 4f^2_{xy}}]/2. \qquad [40]$$

The directions of corresponding eigenvectors, which are always orthogonal to each other since $f_{\alpha\beta}$ is symmetric, are (apart from a factor) $[f_{xy}, \lambda_1 - f_{xx}]^T$ and $[f_{xy}, \lambda_2 - f_{xx}]^T$, respectively. Note that the eigenvalues of $g_{\mu\nu}$ are λ_1^2 and λ_2^2, and the corresponding directions of eigenvectors are the same as those of $f_{\alpha\beta}$ given above. At degenerating points either λ_1 or λ_2 is zero. The corresponding direction (hereby called the *degenerating direction* $\mathbf{t}_D$) becomes

$$\mathbf{t}_D = \begin{bmatrix} f_{xy} \\ -f_{xx} \end{bmatrix}. \qquad [41]$$

The collection of degenerating points forms a curve, which is called the *degenerating curve*. Notice, however, that the degenerating direction is not necessarily related to the direction, either tangent or normal, of the degenerating curve.

PROPOSITION 4. *The directions of all geodesics passing through a degenerating point given by* $f_{xx}f_{yy} - f_{xy}^2 = 0$ *coincide with the degenerating direction.*

Proof. The geodesics obeying Eq. **20** may be written explicitly as

$$\begin{cases} f_x = b_1 s + c_1, \\ f_y = b_2 s + c_2. \end{cases} \quad [42]$$

Substituting s yields a curve in its implicit functional form

$$b_2 f_x - b_1 f_y = c, \quad [43]$$

where $c = c_1 b_2 - c_2 b_1$ is another constant. Note that the normal direction of any curve $F(x, y) = c$ is given by $[F_x, F_y]^T$, and therefore the normal direction of the geodesic $\mathbf{n}_G$ is

$$\mathbf{n}_G = \begin{bmatrix} b_2 f_{xx} - b_1 f_{yx} \\ b_2 f_{xy} - b_1 f_{yy} \end{bmatrix}. \quad [44]$$

The Euclidean inner product of $\mathbf{n}_G$ with $\mathbf{t}_G$ is

$$\begin{aligned} \mathbf{n}_G \cdot \mathbf{t}_D &= f_{xy}(b_2 f_{xx} - b_1 f_{yx}) - f_{xx}(b_2 f_{xy} - b_1 f_{yy}) \\ &= b_1(f_{xx}f_{yy} - f_{yx}f_{xy}) = 0 \end{aligned} \quad [45]$$

at degenerating points for any parameters b_1, b_2 that specify the family of geodesics passing through a particular point. This is equivalent to saying that the tangent direction of any geodesics is along the degenerating direction. q.e.d.

The above proposition indicates that all geodesics approaching the degenerating point must "converge" at the degenerating direction and all geodesics leaving the degenerating point must be confined to the degenerating direction. Here the two-dimensional perceptual space appeared somehow degenerated or "collapsed" into only one direction, since the metric has zero component along the other (orthogonal) direction.

Discussion

Starting from the basic argument about perceptual oneness in rigid motions, we have deduced, without other assumptions, the basic objects of the non-Euclidean perceptual space—the affine connection $\Gamma^{\lambda}_{\alpha\beta}$, the Riemann curvature tensor $R^{\alpha}_{\eta\beta\gamma}$, the geodesics, and the metric tensor $g_{\mu\nu}$. The response of motion detectors resulting from a rigid moving object is *construed* in this perceptual space (by "homunculus") as an intrinsically constant vector field representing the perceptual entity of a single, rigid object. The intrinsic constancy of the detector response may be most easily recognized under the "good" coordinates of the perceptual space—geodesics of any family. The inner product of the detector response V_α and the tangent direction of a geodesic dx^α/ds is constant:

$$V_\alpha \frac{dx^\alpha}{ds} = f_{\alpha\beta}\, v^\beta \frac{dx^\alpha}{ds} = \frac{df_\beta}{ds} v^\beta = b_\beta v^\beta = \text{constant}. \quad [46]$$

In particular, the family of geodesics $[b_1, b_2]^T \propto [-v_2, v_1]^T$ is everywhere orthogonal to the detector output.

Geodesics represent "straight" lines in the perceptual space. The local difference between a geodesic and a physically straight line is an infinitesimal quantity on the order of $o(ds)$. However, at the degenerating points, since all geodesics converge to one direction (the degenerating direction), the local difference between the two may become discernible. Recall the well-known Hering illusion where a pair of physically parallel lines are perceived as being "bent" toward each other in the presence of the background figure of radiating lines or "spokes." This perceptual distortion of angular relationships may be due to the noncoincidence between a "perceptually" straight line and a "physically" straight line at the degenerating curves (the background spokes, presumably). Such visual illusions, according to this explanation, are occasional "artifacts" of the visual system designed for the more important cause of forming the percept of an object defined, in this case, by pure motion information.

The general idea of relating perceptual oneness to the concept of intrinsic constancy under a non-Euclidean geometry may be extended to other visual modalities such as color, depth, etc. The perceptual structure of vision can then be described as a fiber bundle, with visual space as the base manifold, the aforementioned affine connection as the base connection, motion system as the tangent fiber, and all other relevant visual modalities as general fibers. The cross section of the fiber bundle is nothing but a visual scene, an intrinsically constant (parallel) portion of which represents a visual object.

We dedicate this paper to the heroic students and civilians of Beijing who died on June 4, 1989. J.Z. acknowledges the constant encouragement from and enlightening discussions with Drs. R. L. De Valois and K. K. De Valois. J.Z. is supported in part by a research assistantship from Public Health Service Grant EY-00014. S.W. is supported in part by a research assistantship from National Science Foundation Grant PHY-8708447.

1. Hubel, D. H. & Wiesel, T. N. (1959) *J. Physiol. (London)* **148,** 574–591.
2. Hubel, D. H. & Wiesel, T. N. (1962) *J. Physiol. (London)* **160,** 106–154.
3. Hubel, D. H. & Wiesel, T. N. (1968) *J. Physiol. (London)* **195,** 215–243.
4. Maunsell, J. H. R. & Newsome, W. J. (1987) *Annu. Rev. Neurosci.* **10,** 363–401.
5. DeYoe, E. A. & Van Essen, D. C. (1988) *Trends Neurosci.* **11,** 219–226.
6. Livingstone, M. S. & Hubel, D. H. (1988) *Science* **240,** 740–749.
7. Zeki, S. & Shipp, S. (1988) *Nature (London)* **335,** 311–317.
8. Sharpley, R. M. (1990) *Annu. Rev. Psychol.* **41,** 635–658.
9. Hassenstein, B. & Reichardt, W. (1956) *Z. Naturforsch.* **11b,** 513–524.
10. Reichardt, W. (1957) *Z. Naturforsch.* **12b,** 448–457.
11. Poggio, T. & Reichardt, W. (1973) *Kybernetik* **13,** 223–227.
12. van Santen, J. P. H. & Sperling, G. (1984) *J. Opt. Soc. Am. A Opt. Image Sci.* **1,** 451–473.
13. van Santen, J. P. H. & Sperling, G. (1985) *J. Opt. Soc. Am. A Opt. Image Sci.* **2,** 300–321.
14. van Doorn, A. J. & Koenderink, J. J. (1982) *Exp. Brain Res.* **45,** 179–188.
15. van Doorn, A. J. & Koenderink, J. J. (1982) *Exp. Brain Res.* **45,** 189–195.
16. Adelson, E. H. & Bergen, J. R. (1985) *J. Opt. Soc. Am. A Opt. Image Sci.* **2,** 284–299.
17. Watson, A. B. & Ahumada, A. J., Jr. (1985) *J. Opt. Soc. Am. A Opt. Image Sci.* **2,** 322–341.
18. Emerson, R. C., Citron, M. C., Vaughn, W. J. & Klein, S. A. (1987) *J. Neurophysiol.* **58,** 33–65.
19. Reichardt, W. & Schlögl, R. W. (1988) *Biol. Cybern.* **60,** 23–35.
20. Tretiak, O. & Pastor, L. (1984) in *Seventh International Conference on Pattern Recognition* (IEEE Computer Society Press, Silver Spring, MD), Vol. 1, pp. 16–19.
21. Uras, S., Girosi, F., Verri, A. & Torre, V. (1988) *Biol. Cybern.* **60,** 79–87.
22. Adelson, E. H. & Movshon, J. A. (1982) *Nature (London)* **300,** 523–525.
23. Hildreth, E. C. (1984) *The Measurement of Visual Motion* (MIT Press, Cambridge, MA).
24. Reichardt, W., Schlögl, R. W. & Egelhaaf, M. (1988) *Naturwissenschaften* **75,** 313–315.
25. Adler, R., Bazin, M. & Schiffer, M. (1965) *Introduction to General Relativity* (McGraw-Hill, New York).