Perception, 1974, volume 3, pages 63-80

Optical velocity patterns, velocity-sensitive neurons, and space perception: a hypothesis

K Nakayama, J M Loomis¶
Smith-Kettlewell Institute and Department of Visual Sciences, Pacific Medical Center, 2232 Webster Street, San Francisco, California 94115, USA
Received 21 June 1974

Abstract. A hypothesis is put forward of how global patterns of optical flow, as discussed by Gibson, Johansson, and others, could be processed by relatively simple physiological mechanisms. It is suggested that there may exist motion-sensitive cells in the visual system which operate on the optical flow over the retina, and, in so doing, structure the visual field in terms of distinct surfaces that move and/or lie at varying distances from the observer. First, concepts of static and dynamic perspective relative to a sphere centered about the eye are developed, partly on the basis of the work of Gordon. It is pointed out that the velocity flow pattern has a very simple form making it amenable to analysis by relatively low-level mechanisms. Next a higher-order variable of optical flow, the 'convexity', is defined; under the assumption of a rigid environment, convexity is shown to be related to relative depth. It is then postulated that velocity-sensitive cells having center-surround organization could be linked in such a way as to define a higher-order cell, the convexity cell, having functional properties that make it sensitive to the convexity function. The response profile of a layer of such cells would provide an efficient structuring of the visual field in terms of distinct optical surfaces. Relevant evidence is briefly discussed. Lastly, the optical flow patterns corresponding to rotations of the observer are considered. It is shown that the convexity cell is insensitive to rotations and in consequence responds in an invariant fashion to aspects of the optical flow which are related to the surrounding environment.

1 Introduction

Currently in the field of space perception there is a great deal of interest in stereopsis, stemming primarily from the work of Julesz (1971) and from the recent neurophysiological discovery (Barlow *et al.*, 1967) of cortical units which are selectively responsive to retinal disparity. The emerging view is that relative depth information for the entire binocular field is processed in parallel by these disparity-selective cells; thus simultaneous local processing throughout the visual field provides it with global depth organization (Bishop, 1970; Julesz, 1971). In contrast, the so-called monocular empirical cues to depth, like linear perspective and overlay, are thought of as higher-order spatial properties of the structured light array; as such they presumably require more complex mechanisms for extraction, mechanisms that probably develop through visual experience and which are not likely to be understood in terms of currently available neurophysiological concepts.

The view that we have just presented ascribes a special status to stereopsis since it is subserved by a relatively peripheral neurological substrate. It is our belief, however, that possibly as important as stereopsis in the reconstruction of depth from the two-dimensional retinal image is the monocular processing of optical velocities.

It is well known that one-eyed individuals are quite able to drive cars, land airplanes, and perform other equally complex perceptual-motor tasks. Similarly, many species having little or no overlap of the two visual fields are able to maneuver at high speed through branches or brush. We suggest that this kind of performance is made possible primarily by optical velocity information. As we shall show, optical

¶ Present address: Department of Psychology, University of California, Santa Barbara, California 93106, USA.

velocity information, like retinal disparity, could be processed by relatively simple analyzers in the visual system. We believe that retinal disparity and optical flow are likely to be the primary cues for the reconstruction of depth providing the earliest meanings of distance for the developing organism. Although the two seem remarkably similar in terms of their level of processing, we feel that they control different behaviors. Retinal disparity, based on a relatively small interpupillary distance, probably controls behavior which is directed at the near environment; whereas optical velocity information, based on much greater displacements of a single eye, controls more distantly directed behavior. To extend the analogy further, we propose the term 'kineopsis' to correspond to 'stereopsis', the two terms referring to the processes operating on optical flow information and retinal disparity, respectively.

In the first part of the paper we shall develop the meaning of the optical velocity field and show how the velocity field is related to the distance of objects in three-dimensional space. In the second part of the paper we shall consider how a class of velocity-sensitive mechanisms in the visual system might operate on the velocity field, and in so doing structure the visual field in terms of surfaces lying at different distances from the observer.

2 Optical analysis

2.1 *The spherical representation of the optic array*

The significance of the optical flow for space perception has been stated previously, most notably by Gibson (1950, 1966). We add to Gibson's model by asserting the amenability of optical flow to analysis by relatively simple neural mechanisms, in direct analogy to stereopsis.

At this point we would like to clarify the difference in meaning between optical flow and movement parallax. The difference is primarily one of emphasis: movement parallax, one of the classical cues to depth, refers to the relative optical displacement of two environmental points occasioned by movement of the points or of the observer. As it is defined, it emphasizes only a small number of focal objects; in contrast, the optical flow concept emphasizes the totality of optical motions, relative and absolute, induced in the optic array by observer motion (Gibson, 1950). The latter idea is much more suited for thinking of the optical velocity analysis as being performed by relatively simple mechanisms operating in parallel to yield depth information.

We now develop the meanings of the optic array and the optical flow pattern. There are two alternative ways of beginning. The first, put forward by Gibson (1966), proceeds in the following way. Imagine a particular point in space, called a 'station point'. In a well-lit environment the surrounding surfaces are represented at this station point by a complex array of reflected light. If an observer were to position his eye at this point, he would view, for any particular fixational position, a portion of the optic array thus determined. Because the optic array at each station point represents the field of view of an eye that were to assume that position and scan in all directions, Gibson (1966) considers all station points and their corresponding optic arrays as specifying the 'permanent possibilities of vision' (p.191). Although this manner of defining the optic array is attractive in that it reminds us of the enormously intricate structure of reflected light in the open air, it presents difficulties in defining the optical velocity pattern of a moving observer.

Consequently, we choose to develop a meaning of the optic array which is tied to a moving observer's eye rather than to a fixed point in space. This has been rigorously set forth by Gordon (1965). By way of introduction, we mention that the entrance pupil limits the bundle of rays entering the eye from each object point, and that the center of the entrance pupil can be considered the projection center for object space:

similarly, the exit pupil and exit pupil center play the corresponding roles in image space (Fry, 1969; Enoch and Laties, 1971; Enoch and Hope, 1972). Because the light from the surrounding environment is projected essentially through a point (the entrance pupil), only the angular positions of the environmental points are preserved in the projection, all pointwise distance information being lost; that the retina is two-dimensional is irrelevant, for the information is lost prior to reaching the retina in the projection through the entrance pupil. To the extent that the entrance pupil is not really a point there is some slight distance information through blurring and accommodative action of the eye muscles, but we ignore this since it is useful only for very near vision.

Since pointwise distance information is lost in the projection, a stationary observer monocularly viewing the environment might as well be viewing a two-dimensional representation of the environment. Considering this fact, we are interested in representing the array of reflecting surfaces by their projection onto some suitable surface. Of the infinitely many possible surfaces, the plane and the sphere seem most appropriate. The planar representation of the optic array has been most used for representational art and photography and is the form of projection for which the concepts of perspective are most highly developed (Pirenne, 1970). It is not particularly suited, however, for the study of vision, there being two primary disadvantages: the first is that the surface metric corresponding to a given angular extent varies with field position, and the second that at most one half of the environment can be represented by single-image planar representation. A more convenient specification is a central projection of the environment onto a reference sphere of fixed radius. In accordance with what we have said, it is appropriate to consider the reference sphere as being centered on the entrance pupil center.
The optic array being defined here is somewhat of a geometrical abstraction, for, although the entrance pupil in actuality admits a bundle of rays from each object point, only those directed at the pupil center (the 'chief rays') figure in the definition of the optic array by their intersections with the sphere. As it is defined, the optic array is an objective specification of the luminous environment which depends only upon the position of the pupil center; this means that the optic array can be considered effectively independent of the rotational state of the eye. It is true that the entrance pupil (10·5 mm anterior to the center of rotation of the human eye) translates slightly with changes in fixation, but the resulting change in perspective is noticeable only for objects very close to the eye. Therefore we treat the entrance pupil and the center of rotation as being coincident.

Up until now we have said nothing about the image formed by the human eye. Dealing only in object space terms, we have shown that eye rotations permit the stationary observer to scan an essentially unchanging optic array. What remains is to describe what is happening inside the globe.

The effective stimulus for vision is the distribution of light flux falling on the retina. Optically speaking, this distribution is the 'screen image' which is defined as the intersection of the inner surface of the globe with the rays of light projecting from the exit pupil to form the three-dimensional dioptric image. When a stationary observer makes eye rotations, the retina scans a more or less stationary screen image to define, for any one eye position, the so-called 'retinal image'. It is important to mention that in humans the form of the screen image is not identical with the corresponding portion of the optic array, because the globe is only roughly spherical and because the optic array is defined by a projection onto the sphere through its center, whereas the exit pupil, which serves as the projection center for the screen image, is anterior to the approximate center of the globe. It may seem more fitting to define the optic array as reflecting the form of the screen image; but this is unnecessary, if not undesirable, when we recognize that, for a given eye, there is a fixed correspondence, albeit complicated, between the optic array and the totality of screen images (each corresponding to a different fixational position).

By defining the optic array independently of the screen image, we provide a specification of the stimulus which does not depend upon the anatomical or optical characteristics of the particular eye. This permits us to define one and only one optic array for all species, provided only that the eye functions by projection through a point. Using this intermediate construction, we can later take the particular characteristics of the eye in question to derive the effective retinal stimulation.

2.2 *The optical flow pattern under translational motion*

What we have been developing so far is the idea of the optic array, which is the spherical representation of the environment of a stationary observer. Ultimately, we are interested in describing the optical flow pattern of the optic array of a moving observer, as represented on the reference sphere in relation to the three-dimensional environment.

It was Gibson (1950) who first emphasized the ubiquity of optical flow and its importance for the perception of depth. He has alternately dealt with the flow pattern in terms of the plane and the sphere. However, it was Gordon (1965) who most rigorously developed the spherical representation of the optical flow; we take his work as a starting point.

Consider an observer moving toward some point D. Each visible point in the environment can be represented by its projection image on the surface of the sphere. If we know the Cartesian coordinates of each point, it is a simple matter to specify each image either in terms of longitude and latitude (Gordon, 1965) or in terms of meridian and eccentricity (Whiteside and Samuel, 1970) with the horizontal plane of the head and the direction of D serving to define the axes. The totality of position vectors (coordinates) representing the images of environmental points defines the 'instantaneous positional field' (Gordon, 1965) at each point in time. For the sake of absolute clarity, we bring attention to the fact that each image point on the sphere corresponds to a unique environmental point; this follows from the obvious fact that the nearest reflecting point in a particular direction from the eye occludes the light from all other more distant points in that direction.

Before introducing dynamic perspective and the meaning of the instantaneous velocity field, it will be helpful to develop the idea of static perspective in the instantaneous positional field.

Imagine a family of parallel lines lying in Cartesian space. Define as the 'ancillary line' (Pirenne, 1970) that particular line in the family which passes through the center of the sphere (the pupil center). The two intersections of the ancillary line with the surface of the sphere determine the 'vanishing points'. The projected images of all the parallels, except the ancillary line, are great circles passing through these two vanishing points.

We now define the optical flow pattern or, more precisely, the 'instantaneous velocity field' (Gordon, 1965). As our imaginary observer moves in space toward the point D, every visible point (assumed fixed) in the environment has a determinate angular velocity with respect to the entrance pupil (the center of the reference sphere). The instantaneous velocity field associates with each image point of the instantaneous positional field the angular velocity of the corresponding environmental point. An illustration of the spherical field for a subject flying over a flat surface is shown in figure 1.

In order to develop a sense of the velocity field, we begin by considering the family of parallel lines in the direction of the observer's motion (toward D). Recall that the images of all these parallels are great circles (meridians) passing through the two points where the ancillary line (in this case, the direction of motion) intersects the sphere. Now consider any stationary point in space; since it necessarily lies on one of the parallels, the image of that point will trace, over successive positional fields, the meridian corresponding to that parallel line provided that the observer moves with constant direction. More importantly, the direction of the angular velocity will always be tangential to that same meridian. Since every point in space lies on one of the infinitely many parallels, it follows that the *directions* of flow are always along the meridians, provided that the environment is rigid and that there is no instantaneous change of direction of the observer's motion.

With the aid of the angular measures of meridian (α) and eccentricity (β), derivation of the angular velocities is extremely simple (Whiteside and Samuel, 1970). From the above considerations, the velocity *directions* are determined by the position vectors (assuming, once again, a rigid environment and constant direction of motion) so that we need only derive the velocity *magnitudes*. Referring to figure 2, consider the observer O moving toward D with velocity V, the stationary point Q has the spherical coordinates of α_Q, β_Q and distance S (lost in the projection) where these are defined with respect to O and the direction of D. The value of α (meridian) does not come into the derivation of velocity magnitude since velocity magnitude is independent of α, all other things being equal. Since the observer is moving with velocity V, the

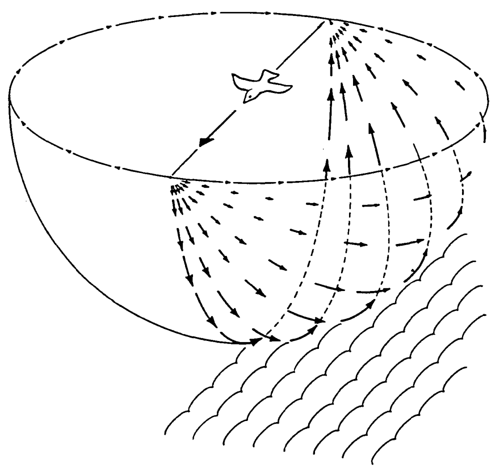

Figure 1. Optical flow pattern for a bird flying over the ground, showing that the flow is directed along great circles on the reference sphere (from Gibson, 1966; reproduced with permission of the author).

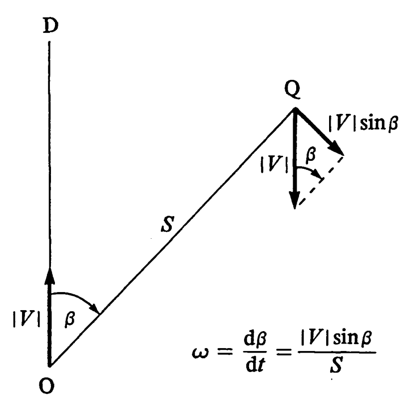

Figure 2. Situation defined by observer O moving with velocity V toward the point D. Q represents any environmental point. V is the magnitude of velocity of point Q relative to O, and S is the distance from O to Q.

stationary point Q has the relative velocity V with respect to O. The component of velocity which is normal to the direction of Q has the magnitude $|V|\sin\beta_Q$ so that the angular velocity with respect to O has, by definition, the magnitude of

$$|\omega| = \frac{|V|\sin\beta_Q}{S} \tag{1}$$

(see the Appendix for a more general specification of the velocity field). It follows that for any given position in the instantaneous positional field the angular velocity is inversely proportional to the distance of the corresponding environmental point. Thus, although pointwise distance information is lost in the projection through the entrance pupil, absolute distance information can be recovered, in principle, from the absolute angular velocities when $|V|$ is known.

As it is defined, the concept of optical flow is highly abstract in that each and every environmental point projecting to the eye has associated with it a determinate angular velocity. In actuality, there will be situations where an environmental point will be indistinguishable from other points in surrounding angular directions; no processing mechanism can respond to a given environmental point unless the corresponding light flux for that point is discriminable from those of neighboring points in luminance or color. However, our discussion is weakened little by this consideration, for much of the visual environment has texture which is perceptible.

We can briefly summarize the properties of the velocity field for these conditions by saying that, whereas the velocity *directions* are strictly determined by field position, the velocity *magnitudes* depend upon both field position and environmental distance. Thus two adjacent image points in the positional field corresponding to environmental points at different distances have different velocity magnitudes but a common direction. A planar representation of the velocity field for an observer approaching a single rectangular screen while moving above the infinite surface plane is depicted in figure 3a. Note the simplicity of the pattern of velocity directions; this simplicity conflicts somewhat with our intuitive appraisal that the directions of optical flow are complicated, seeming to depend upon the nearness of the different surfaces.

As we did with static perspective, we must now relate the objective velocity field to the optical flow pattern on the retina. For the time being, we shall simplify the discussion by assuming that the observer is translating through the environment without eye, head, or body rotations. This being the case, the optical flow in relation to the retina mirrors the velocity field to the extent that the retina is spherical and concentric about the exit pupil center. As we have indicated earlier, this is not true for humans, there being a consequent transformation of the flow pattern. For our purposes, however, this slight transformation is not important; thus, we consider a 'conceptual' retina which is spherical and centered on the exit pupil. Rotations will be dealt with later using this 'conceptual' retina.

3 An ideal observer

Before going on to the discussion of mechanisms, we consider in detail the information available to an ideal observer moving with constant direction through a fixed environment.

From preceding discussion, the absolute angular velocity of any point Q in the optic array has the magnitude

$$|\omega| = \frac{|V|\sin\beta_Q}{S}\ ; \tag{2}$$

its direction is along one of the great circles passing through the point on the sphere representing the direction of locomotion. By transposing equation (2) the absolute distance of any environmental point can be expressed as

$$S = \frac{|V|\sin\beta_Q}{|\omega|} . \tag{3}$$

Thus an ideal observer knowing his own speed (velocity magnitude) can 'know' the distance S, since β and $|\omega|$ are given in the optic array.

Where $|V|$ is not known, or where only relative angular velocities are registered, the following expression is more appropriate:

$$\frac{S_1}{S_2} = \frac{|\omega_2|\sin\beta_1}{|\omega_1|\sin\beta_2} . \tag{4}$$

This expression relates the distance of one point in the field to that of any other point. Clearly, if the absolute distance of just one environmental point is known, all others can be determined by the ideal observer. Such a point might, for example lie on the ground surface directly below a terrestial organism, for this distance is essentially constant.

The purpose in discussing the ideal observer is only to demonstrate that, on the assumption of a rigid environment, optical velocity information can in principle be used by the observer to reconstruct the depth dimension lost in the projection through the pupil. In our following discussion of mechanisms, however, we do not take the position that optical velocities are mapped point-for-point by the nervous system into perceived distances; to do so would be to ignore the fact that space perception is an exceedingly complex process, depending on many interacting levels of neural organization (Ittelson, 1960; Gogel, 1973a, 1973b). This complexity stems in part from the fact that for any given static or dynamic optic array, ambiguity is always present when the assumption that the environment is rigid is removed (Johansson, 1970; Eriksson, 1973); this follows from the fact there is a multiplicity of 'equivalent configurations' (Ittelson, 1960) of depth and velocity that can produce a given velocity pattern of an environmental object. Thus we postulate only that there are neural mechanisms which respond to optical flow and in so doing structure the visual field in terms of distinct environmental surfaces; how these distinct surfaces are perceived in terms of their velocity and depth depends on high-level organizational factors. The following section presents the hypothesis in detail.

4 Mechanisms for extraction of depth information from the velocity field

Under normal circumstances when an organism moves through the environment there is an enormous amount of redundancy in the pattern of visual stimulation, both spatially and temporally (Attneave, 1954; Barlow, 1961). Therefore much of the velocity-based depth information available to the organism could be coded in a much more efficient manner if the visual system were sensitive to some higher-order variable of the retinal flow rather than to the individual angular velocities. In other contexts Barlow (1969) has suggested that this is precisely what visual 'feature analyzers' do when they respond to a particular stimulus configuration: they are more efficiently coding the stimulus by taking advantage of the redundancy in the visual world.

The major thesis of this paper is that there is a higher-order variable of the optical flow pattern which is both efficient and highly informative about the organization of the surrounding environment in terms of depth, and that there are neural structures in the visual system which code this variable. Our thesis closely parallels the current thinking on the mechanisms which operate on luminance distributions to provide an awareness of surface reflectance (Cornsweet, 1970; Land and McCann, 1971; Ratliff, 1965).

One possible candidate for such a higher-order variable is the 'gradient' of velocity originally discussed by Gibson (1950), though not in the same context. He suggested that the relative velocities between points in the optic array are highly informative about the three-dimensional environment, since large relative velocities of adjacent points are observed in the optic array at the edges of surfaces lying at different distances; furthermore, relative velocities are invariant with respect to eye movements. The difficulty with simple relative velocity, expressed as a difference, ratio, or derivative is that it contrasts the velocity magnitude of one point with that of others only in one direction. We would like to find a function which is sensitive to discontinuities in velocities independently of their directions on the sphere.

4.1 *The convexity function—a center-surround model*

As our candidate for the higher-order variable we propose a function which we call the 'convexity' function. We will first define it and then clarify its meaning by referring to a possible physiological structure which could operate on the retinal flow as a convexity function.

The specific convexity function we have in mind is defined for each angular position (α, β) in the field as

$$C(\alpha,\beta) = \sum_i \left[\int_C V_i - k \int_S V_i \right] . \qquad (5)$$

What this means is that the convexity function C assigns to each (α, β) a scalar value which is determined by the optical flow over the center region C and the concentric surrounding region S, both centered on (α, β). Specifically, it compares the optical flow over center and surround in different orientations, designated by subscript i. Accordingly, V_i refers to the component of optical flow in the direction determined by the value of i. The constant k takes into account the different areas of C and S so that, when the total flow in a given direction is equal for the two regions, the contribution of that particular orientation to the overall value of C is 0.

We chose a specific function for convexity to simplify the concept. However a more general convexity function could be expressed as:

$$C(\alpha,\beta) = \psi_i \left[\mathrm{F}_i \left(\int_C V_i , \int_S V_i \right) \right] . \qquad (6)$$

Here the comparison for each orientation between center and surround is not necessarily subtractive nor are the contributions of each orientation to the overall value of C necessarily cumulated additively. We have theoretical reasons for preferring the subtractive assumption as will become clear later.

As it is defined, the convexity function operates locally throughout the positional field. In effect, it is a function which, when applied to the optical flow of an observer moving through space, immediately gives the edges of all surfaces lying at different depths because it is sensitive to discontinuities in optical flow. Thus, it provides global organization of the field in terms of depth through purely local computations. Land and McCann (1971) arrived at a similar result for brightness by applying an analogous function.

If the areas C and S are sufficiently small, the scalar field resulting from the application of this function to the velocity field is relatively indifferent to slight distortions of the field, such as the change in going from the spherical velocity field to the transformed optical flow on the retina. Accordingly, when the globe is stationary (relative to the head), we can speak in terms of the objective velocity field or, alternatively, in terms of the retinal flow pattern.

One convenient property of the convexity function is that the resulting scalar field will portray the edges of objects independently of the observer's motion. In figure 3a we see the velocity field (planar representation) for an observer who is approaching a point beyond and to the side of a vertically oriented screen and in figure 3b for an observer who is passing off on the right. The resulting scalar fields after application of the convexity function are seen in figures 3c and 3d, showing that the convexity function picks up invariant aspects of the environment. The same qualitative results also obtain for different speeds of approach. Although the velocity flow pattern differs depending on the direction and speed of observer motion, the relative output of the convexity function is the same. As we shall see in the discussion on rotations, the convexity function continues to extract the same invariant features when rotations of any sort are present [1].

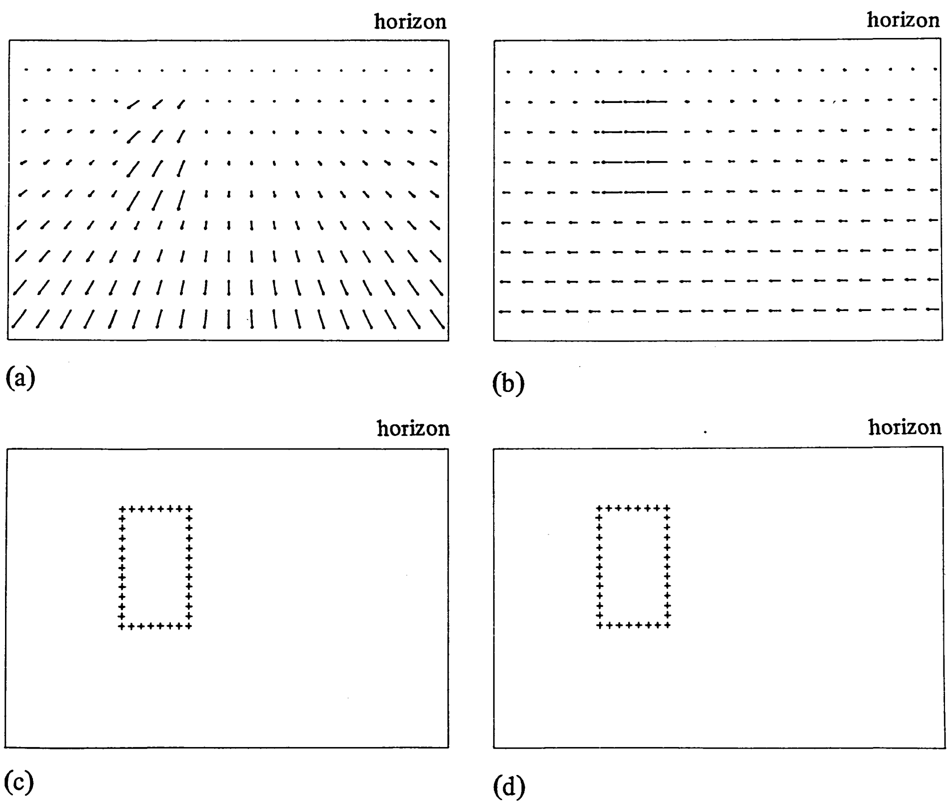

Figure 3. Planar representations of the velocity field of an observer who is (a) approaching a vertical rectangular screen which is off to the side and above the surface plane, and (b) passing a vertical rectangular screen in a direction parallel to the screen; and spatial response profiles (c) and (d) of a layer of the hypothesized 'convexity' cells for the motion in (a) and (b) respectively.

[1] We have chosen the subtractive assumption, equation (5), in defining the convexity, as it has properties which make it invariant with respect to head and eye rotations. Instead, we could have selected a ratio function which takes the ratio of the velocity of the center to that of the surround. This would be equivalent to specifying

$$C(\alpha,\beta) = \sum_i \left(\int_C \log V_i - \int_S \log V_i \right) .$$

An interesting property of this ratio function is that the convexity will remain absolutely invariant with respect to changes in velocity of the observer or field position, whereas the convexity function detailed in equation (5) remains only relatively invariant with respect to different velocities and differing portions of the visual field.

4.2 A physiological model: center–surround cells organized with respect to velocity
There would be little point in identifying a higher-order function of the flow pattern were there not mechanisms for responding to it. We propose that, by virtue of a simple hierarchical connectivity of velocity-sensitive neurons, there is a class of cells at some higher level of the visual pathway having the functional properties of the convexity function. Thus the spatial profile of the outputs of this class of cells reflects the edges of surfaces that are themselves moving or are stationary in an environment through which the observer is moving.

Specifically, we propose that velocity-sensitive cells, so prominent in the visual systems of most species, are organized in the fashion shown in figure 4. Since many velocity-sensitive cells are preferentially sensitive to velocities of a particular direction, we have assumed that velocity cells in the center and surround regions of the convexity cell having a common velocity preference are linked antagonistically; consequently, if the components of optical flow in that direction are different in magnitude, a difference is registered. The overall response of a given convexity cell is the sum of all differences in each orientation at the position in question. This 'wiring diagram' expresses qualitatively the subtractive and additive functions of equation (5); as such, the convexity cell is particularly sensitive to discontinuities of optical flow across the receptive field of the cell, independently of direction. Furthermore, the response of the cell is monotonically related to the size of the discontinuity, providing useful information for the reconstruction of depth.

It should be noted that each velocity-sensitive cell registers the velocity flow component in its direction of preference. This means that each contributing center–surround arrangement responds to the difference in flow components in its direction of preference. The significance of this will be made clear in the discussion of rotations.

As we stated earlier, we are not proposing that the responses of these cells are directly interpreted as relative egocentric distances; rather, we suggest that the response profile of a layer of such cells merely provides an efficient coding of the environment in terms of distinct surfaces. Higher-level interpretative mechanisms would then use this efficient coding together with kinesthetic and other optical information to construct a perceptual field of stationary and moving objects at various distances from the observer. What makes the hypothesis attractive is that a level of global organization of the optic array arrived at through purely local

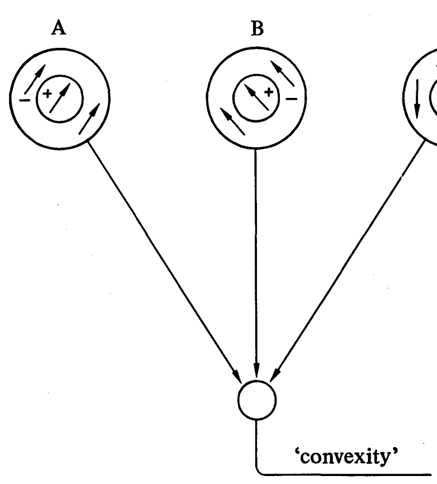

Figure 4. Wiring diagram of the hypothetical 'convexity' cell. Directionally-selective velocity cells with center–surround organization of their receptive fields converge onto a higher-order convexity cell. The receptive fields of the subordinate cells (A, B, C) are all centered on the same retinal locus.

computations is achieved independently of other higher-order variables; this is consistent with the fact that velocity-sensitive cells have been found that are relatively independent of luminance, contrast, and stimulus shape (Barlow and Hill, 1963). That velocity-sensitive cells must play such a role seems obvious upon considering that optical motion resulting from observer movement is more pervasive than optical motion resulting from distal motion in the environment.

5 Evidence for the theory

The point that optical flow information is useful in the perception of depth is already widely appreciated; foremost among researchers, Gibson has brought this idea into prominence with his seminal work (1950) and with later research by himself and his colleagues (Gibson *et al.*, 1955; Gibson, 1957, 1966). Equally important are the investigations of Braunstein (1962, 1966, 1968, 1972) and of Johansson and his associates (Börjesson and von Hofsten, 1972, 1973; Johansson, 1964, 1973; Marmolin, 1973a, 1973b) which have explored the relationship between optical motions and the perception of space, distal motion, and distal events. Where we add to these other writers is in our thesis that there must be relatively peripheral physiological mechanisms for analyzing optical flow information, which is one of the more ubiquitous and largely invariant aspects of visual stimulation.

Neurophysiological recordings from cells showing a center–surround organization with respect to velocity and from cells showing the higher-order property of convexity would provide the most direct evidence for our theory. Although we know of no systematic efforts to find such cells, there are a number of incidental observations which suggest that cells like this may exist. Velocity-sensitive neurons have inhibitory surrounds which are sensitive to stimulus movement (Barlow and Levick, 1965; Sterling and Wickelgren, 1969). For example, Sterling and Wickelgren observed that the center response of a cat superior colliculus cell is inhibited by motion in the surround and that the magnitude of surround inhibition is dependent on the direction of the moving stimulus in the surround, there being more inhibition when the stimulus is moving in the same direction than when it is moving in the opposite direction. Furthermore, Bridgeman (1972, 1973) has reported the existence of cells in the monkey visual cortex which are preferentially activated when a stimulus is moving relative to its surround rather than when the stimulus is moving with respect to the retina.

A line of indirect evidence supporting the theory comes from a study concerned, not with depth perception, but with motion perception (Loomis and Nakayama, 1973). The study describes a velocity illusion which is analogous to a case of brightness contrast. Briefly, if two targets are moving horizontally across a frontoparallel plane with the same angular velocity, and are viewed against a low-density field of moving dots having a horizontal gradient of velocity, the one target which is moving relatively faster than its immediate surround appears to have a higher absolute velocity, even though the two targets maintain a constant separation. The dependence of the perceived velocity of a target on its angular velocity in relation to its immediate surround was interpreted much in the way that the predominance of local luminance ratios (or differences) in determining perceived brightness has been interpreted (Cornsweet, 1970; Ratliff, 1965). It was proposed then that there are motion sensitive cells with antagonistic center–surround organization that respond to local velocity differences or ratios, much in the way that a convexity cell would function. Whether these neural responses are interpreted as differences in velocity and/or in depth depends upon higher-level organizational processes.

6 Rotations

Earlier we developed an objective specification of the optical flow pattern of an observer moving with constant direction (pure translational motion) through the environment. In relating this objective velocity field to the flow pattern on the retina for a stationary eye, we said that there are some global differences, because the retina is not truly spherical and because the projection center is anterior to the center of curvature. However, local properties of the flow pattern, such as the convexity, are preserved. This being the case, we simplified the presentation by considering a spherical 'conceptual retina' with the projection center of image space being at its center. What we have yet to consider are the effects of rotations, either eye rotations or rotations of the observer in space.

The two classes of rotations differ only in relation to the conceptual specification of the objective optic array, for, as we defined it, eye rotations would have no effect. However, if we consider rotations in relation to the conceptual retina, all rotations are equivalent.

Accordingly, we consider first the optical flow on the conceptual retina of a pure rotation about some axis of the sphere. A formal derivation of the rotational field in term of meridian (α) and eccentricity (β), to make it compatible with the specification of the translational field, is given in the Appendix. However, because the directions of flow during a rotation are along the latitudes defined with respect to the axis of rotation, the rotational field is more simply expressed in terms of longitude and latitude just as the translational field is more easily specified in terms of meridian and eccentricity. Briefly then, the optical flow is always directed along the latitude (θ) and the magnitude is $|\omega_R| = |V_R|\sin\theta$ (constant as a function of longitude), where V_R is the rotational velocity and θ the latitude. The important fact here is that the angular velocity of any environmental point is independent of its distance from the observer. Accordingly, for any sufficiently local region of the sphere, the optical flow is uniform. It follows then that convexity is zero everywhere and that no convexity cells respond.

We now consider the velocity field on the conceptual retina for a combined rotation and translation; this is the equivalent optical situation when the observer is moving in curvilinear fashion (see Appendix) or is making pursuit eye movements while translating, such as fixating a target off to the side. Under these conditions the velocity field at each point on the retina is the vector sum of the translational and rotational fields; for a formal derivation see the Appendix. Unlike either in a pure rotational field or in a pure translational field, the directions of flow of a combined field no longer have local uniformity. However, because the rotational flow component is everywhere locally uniform (both in magnitude and direction), and because each constituent cell of the convexity analyzer registers the optical flow (component) in its direction of sensitivity and contrasts the center activity with that of the surround, the response of the convexity cell is essentially invariant under all added rotational fields. Thus, the layer of convexity cells outlines the edges of environmental surfaces or moving objects independently of eye rotations or rotations of the body relative to fixed environmental directions.

7 Summary

We have developed a specification of the instantaneous flow pattern of an observer moving through a rigid environment. We have also suggested that a layer of motion-sensitive cells with appropriate functional properties could serve to outline the edges of surfaces lying at different depths from the observer or moving through the environment. This hypothesis provides a way of understanding how optical motion can lend a primitive structure to the visual field, a structure which is then interpreted by higher-order mechanisms in terms of motion and depth.

Acknowledgements
This investigation was supported by NIH research grants Nos 5 TO1 EY-00037 and 3 PO1 EY-00299 from the National Eye Institute and by the Smith-Kettlewell Eye Research Foundation.
We appreciate the helpful advice of Albert B. Alden regarding the analyses in the Appendix.

References
Attneave, F., 1954, "Some informational aspects of visual perception", *Psychological Review*, **61**, 183-193.

Barlow, H. B., 1961, "Possible principles underlying the transformations of sensory messages" in *Sensory Communication*, Ed. W. A. Rosenblith (Wiley, New York), pp.217-234.

Barlow, H. B., 1969, "Trigger features, adaptation and economy of impulses" in *Information Processing in the Nervous System*, Ed. K. N. Leibovic (Springer, New York), pp.209-226.

Barlow, H. B., Blakemore, C., Pettigrew, J. D., 1967, "The neural mechanism of binocular depth discrimination", *Journal of Physiology*, **193**, 327-342.

Barlow, H. B., Hill, R. M., 1963, "Selective sensitivity to direction of movement in ganglion cells of the rabbit retina", *Science*, **139**, 412-414.

Barlow, H. B., Levick, W. R., 1965, "The mechanism of directionally selective units in rabbit's retina", *Journal of Physiology*, **178**, 477-504.

Bishop, P. O., 1970, "Beginning of form vision and binocular depth discrimination in cortex" in *The Neurosciences: Second Study Program*, Ed. F. O. Schmitt (Rockefeller University Press, New York), pp.471-485.

Börjesson, E., von Hofsten, C., 1972, "Spatial determinants of depth perception in two-dot motion patterns", *Perception and Psychophysics*, **11**, 263-268.

Börjesson, E., von Hofsten, C., 1973, "Visual perception of motion in depth: application of a vector model to three-dot motion patterns", *Perception and Psychophysics*, **13**, 169-179.

Braunstein, M. L., 1962, "The perception of depth through motion", *Psychological Bulletin*, **59**, 422-433.

Braunstein, M. L., 1966, "Sensitivity of the observer to transformations of the visual field", *Journal of Experimental Psychology*, **72**, 683-689.

Braunstein, M. L., 1968, "Motion and texture as sources of slant information", *Journal of Experimental Psychology*, **78**, 247-253.

Braunstein, M. L., 1972, "Perception of rotation in depth", *Psychological Review*, **79**, 510-524.

Bridgeman, B., 1972, "Visual receptive fields to absolute and relative motion during tracking", *Science*, **178**, 1106-1108.

Bridgeman, B., 1973, "Receptive fields in single cells of monkey visual cortex during visual tracking", *International Journal of Neuroscience*, **6**, 141-152.

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

Enoch, J. M., Hope, G. M., 1972, "An analysis of retinal receptor orientation III. Results of initial psychophysical tests", *Investigative Ophthalmology*, **11**, 765-782.

Enoch, J. M., Laties, A. M., 1971, "An analysis of retinal receptor orientation II. Prediction for psychophysical tests", *Investigative Ophthalmology*, **10**, 959-970.

Eriksson, E. S., 1973, "Distance perception and the ambiguity of visual stimulation: a theoretical note", *Perception and Psychophysics*, **13**, 379-381.

Fry, G. A., 1969, *Geometrical Optics* (Chilton Book Company, New York).

Gibson, J. J., 1950, *The Perception of the Visual World* (Houghton Mifflin, Boston).

Gibson, J. J., 1957, "Optical motions and transformations as stimuli for visual perception", *Psychological Review*, **64**, 288-295.

Gibson, J. J., 1966, *The Senses Considered as Perceptual Systems* (Houghton Mifflin, Boston).

Gibson, J. J., Olum, P., Rosenblatt, F., 1955, "Parallax and perspective during aircraft landings", *American Journal of Psychology*, **68**, 372-385.

Gogel, W. C., 1973a, "The organization of perceived space I. Perceptual interactions", *Psychologische Forschung*, **36**, 195-221.

Gogel, W. C., 1973b, "The organization of perceived space II. Consequences of perceptual interactions", *Psychologische Forschung*, **36**, 223-247.

Gordon, D. A., 1965, "Static and dynamic visual fields in human space perception", *Journal of the Optical Society of America*, **55**, 1296-1303.

Ittelson, W. H., 1960, *Visual Space Perception* (Springer, New York).

Johansson, G., 1964, "Perception of motion and changing form", *Scandinavian Journal of Psychology*, **5**, 181-208.

Johansson, G., 1970, "On theories for visual space perception", *Scandinavian Journal of Psychology*, **11**, 67-74.
Johansson, G., 1973, "Visual perception of biological motion and a model for its analysis", *Perception and Psychophysics*, **14**, 201-211.
Julesz, B., 1971, *Foundations of Cyclopean Perception* (University of Chicago Press, Chicago).
Land, E. H., McCann, J. J., 1971, "Lightness and retinex theory", *Journal of the Optical Society of America*, **61**, 1-11.
Loomis, J. M., Nakayama, K., 1973, "A velocity analogue of brightness contrast", *Perception*, **2**, 425-428.
Marmolin, H., 1973a, "Visually perceived motion in depth resulting from proximal changes I", *Perception and Psychophysics*, **14**, 133-142.
Marmolin, H., 1973b, "Visually perceived motion in depth resulting from proximal changes II", *Perception and Psychophysics*, **14**, 143-148.
Pirenne, M. H., 1970, *Optics, Painting and Photography* (Cambridge University Press, Cambridge).
Ratliff, F., 1965, *Mach Bands* (Holden Day, San Francisco).
Sterling, P., Wickelgren, B. G., 1969, "Visual receptive fields in the superior colliculus of the cat", *Journal of Neurophysiology*, **32**, 1-15.
Whiteside, T. C. D., Samuel, G. D., 1970, "Blur zone", *Nature*, **225**, 94-95.

APPENDIX

Here we derive the instantaneous velocity field with respect to the objective sphere for a variety of observer motions and then indicate the nature of the flow pattern with respect to the mobile conceptual retina. The general problem is to specify for each position on the sphere the angular velocity of the corresponding environmental point. The velocity field then, as Gordon (1965) noted, is really a positional velocity field, for it is a field of velocity vectors associated with a field of position vectors. It is helpful in the following derivations to keep in mind the two distinct vector fields.

We first derive the velocity field on the objective sphere for three classes of motion: (1) pure translation through the environment, (2) pure rotation and (3) translation plus rotation, showing also that the instantaneous velocity field of curvilinear motion (changing direction) is equivalent to the latter case. We then indicate the nature of the velocity field relative to the conceptual retina of an observer making pursuit eye rotations while undergoing any combination of rotational and translational motion.

Case 1. Assume that the coordinates of the objective sphere are defined with respect to the orbit, with the primary position of the eye defining the X axis and the frontal plane of the head determining Y and Z. If the observer's head is erect and facing the direction of motion, these axes coincide with the axes assumed in the text (the direction of motion and the surface plane). This slight complication allows us to deal with the next case where there is no translational motion.

The most convenient way of conceptualizing each angular velocity vector is as a vector tangent to the sphere of unit radius. Thus we need to find the direction and magnitude of each velocity vector $\vec{\omega}_T(\alpha, \beta)$ for each environmental point Q at (α, β, S) in the direction $\vec{Q} = (\alpha, \beta, 1)$ as the observer moves in the direction $\vec{O}$ (origin vector along X) with velocity V_T (figure A1). In the figure, both spherical and Cartesian coordinate systems are shown, for we find it simplest to specify each position vector (α, β) in terms of meridian and eccentricity ($0 \leqslant \alpha \leqslant 2\pi$, $0 \leqslant \beta \leqslant \pi$), while expressing each velocity vector in terms of the *lengths* of its components along

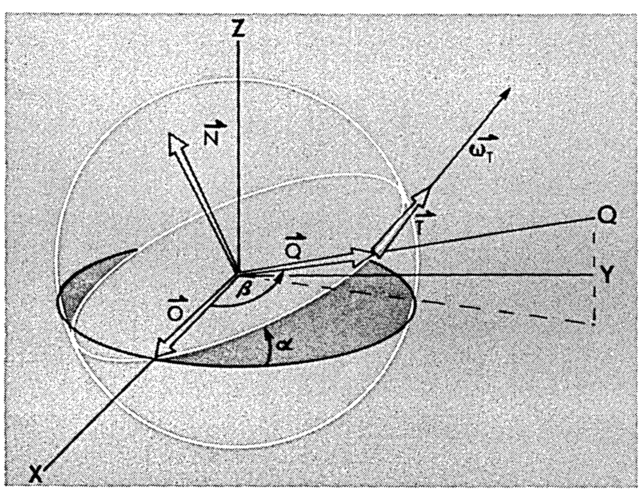

Figure A1. The translational velocity field. Pictured are spherical and Cartesian coordinate systems centered on the entrance pupil. Any environmental point Q is represented by its angular coordinates of meridian (α) and eccentricity (β) and its distance S. $\vec{O}$ is the unit origin vector in the direction of motion (along the X axis), $\vec{Q}$ is a unit vector in the direction of Q, $\vec{N}$ is a unit vector normal to $\vec{O}$ and $\vec{Q}$, and $\vec{\omega}_T$ represents the instantaneous angular velocity of point Q in the direction of the unit vector $\vec{T}$.

the axes of X, Y, Z space. The origin of each velocity vector then is the corresponding position vector $(\alpha, \beta, 1)$ in spherical coordinates or $(\cos\beta, \sin\beta\cos\alpha, \sin\beta\sin\alpha)$ in Cartesian coordinates.

From the text, the vector $\vec{\omega}_T(\alpha, \beta)$ corresponding to Q has a magnitude

$$\frac{|V_T|\sin\beta}{S} \,.$$

Also from the text, its direction is along the meridian defined by the direction of motion; this direction can be expressed by the unit vector $\vec{T}$ where

$$\vec{T} = \frac{\vec{N} \times \vec{Q}}{|\vec{N} \times \vec{Q}|} = \vec{N} \times \vec{Q}$$

where $\vec{N}$ is the position vector (a unit vector) normal to $\vec{O}$ and $\vec{Q}$. Because

$$\vec{N} = \frac{\vec{O} \times \vec{Q}}{|\vec{O} \times \vec{Q}|} = \frac{\vec{O} \times \vec{Q}}{\sin\beta} \,,$$

we have

$$\vec{T} = \frac{1}{\sin\beta}(\vec{O} \times \vec{Q}) \times \vec{Q} \,.$$

When these vectors are expressed in Cartesian coordinates, the angular velocity vector for position (α, β) becomes

$$\omega_T(\alpha, \beta) = \frac{|V_T|\sin\beta}{S\sin\beta}\left[\begin{pmatrix}1\\0\\0\end{pmatrix} \times \begin{pmatrix}\cos\beta\\ \sin\beta\cos\alpha\\ \sin\beta\sin\alpha\end{pmatrix}\right] \times \begin{pmatrix}\cos\beta\\ \sin\beta\cos\alpha\\ \sin\beta\sin\alpha\end{pmatrix} = \frac{|V_T|}{S}\begin{pmatrix}-\sin^2\beta\\ \cos\beta\sin\beta\cos\alpha\\ \cos\beta\sin\beta\sin\alpha\end{pmatrix} . \tag{A1}$$

Case 2. We now derive the velocity field for a pure rotation. Because the instantaneous flow pattern of an observer rotating simultaneously about two axes (one system within the other) is equivalent to that of a single rotation, we need only treat the

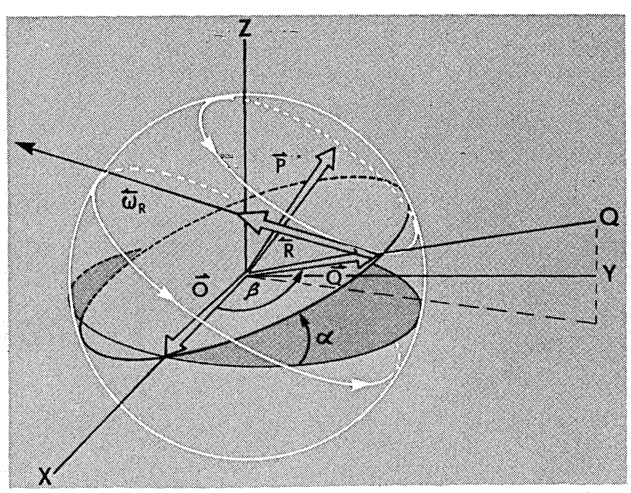

Figure A2. The rotational velocity field. Pictured are spherical and Cartesian coordinate systems centered on the entrance pupil. Any environmental point Q is represented by its angular coordinates of meridian (α) and eccentricity (β). $\vec{O}$ is a unit vector along the X axis, $\vec{Q}$ is the position vector of point Q, $\vec{P}$ is the unit vector defining the positive pole of the axis of rotation, and $\vec{\omega}_R$ is the rotational velocity vector of point Q in the direction of the unit vector, $\vec{R}$.

latter. Consider an observer at the origin rotating (relative to the environment) at V_R rad s^{-1} about a single axis specified by the position vector $\vec{P} = (\alpha_P, \beta_P, 1)$ of its positive pole (figure A2). Because the image of any environmental point will trace an arc length on the rotating sphere which is proportional both to V_R and $\sin\theta$, where θ is the angular separation of the pole and the point Q, the magnitude of $\vec{\omega}_R$ is $|V_R|\sin\theta$. The direction of the flow vector is given by the unit vector $\vec{R}$ which is normal to both $\vec{P}$ and $\vec{Q}$. Thus

$$\vec{R} = \frac{\vec{Q}\times\vec{P}}{|\vec{Q}\times\vec{P}|} = \frac{\vec{Q}\times\vec{P}}{\sin\theta} \; .$$

Expressing the position vectors in Cartesian coordinates, we obtain the angular velocity for Q in the form

$$\vec{\omega}(\alpha,\beta) = \frac{|V_R|\sin\theta}{\sin\theta}\begin{pmatrix}\cos\beta \\ \sin\beta\cos\alpha \\ \sin\beta\sin\alpha\end{pmatrix} \times \begin{pmatrix}\cos\beta_P \\ \sin\beta_P\cos\alpha_P \\ \sin\beta_P\sin\alpha_P\end{pmatrix}$$

$$= |V_R|\begin{pmatrix}\sin\beta\cos\alpha\sin\beta_P\sin\alpha_P - \sin\beta\sin\alpha\sin\beta_P\cos\alpha_P \\ \sin\beta\sin\alpha\cos\beta_P - \cos\beta\sin\beta_P\sin\alpha_P \\ \cos\beta\sin\beta_P\cos\alpha_P - \sin\beta\cos\alpha\cos\beta_P\end{pmatrix} \; . \qquad \text{(A2)}$$

Case 3. If the observer is simultaneously translating and rotating in space, the instantaneous optical flow at each point on the sphere is simply the vector sum of the translational and rotational velocity vectors. Because $\vec{\omega}_R$ and $\vec{\omega}_T$ are both tangent to the sphere by virtue of being normal to the position vector $\vec{Q}$, the resultant is also normal to $\vec{Q}$ and is therefore an angular velocity vector. Thus

$$\vec{\omega}(\alpha,\beta) = \vec{\omega}_T(\alpha,\beta) + \vec{\omega}_R(\alpha,\beta) \; . \qquad \text{(A3)}$$

Now consider an observer who is moving through space with changing direction and with velocity V_T; instantaneously, he can be considered to be moving with radius of curvature r about some distant point C having the coordinates (α_C, β_C, r) and located on the sphere by the position vector $\vec{C} = (\alpha_C, \beta_C, 1)$. Even though the observer's motion cannot be decomposed into a pure rotation and a pure translation, his instantaneous velocity field at each point in space is equivalent to that of an observer rotating and translating at the same point. In particular, the velocity field here is the sum of the translational field for velocity V_T and of the rotational field for a rotational velocity of V_T/r (rad s^{-1}) with axis of rotation perpendicular to the instantaneous direction of motion $\vec{O}$ and the direction of C, given by $\vec{C}$.

Thus for the point Q the rotational velocity is

$$\vec{\omega}_R(\alpha,\beta) = \frac{|V_T|}{r}(\vec{Q}\times\vec{P}) = \frac{|V_T|}{r}\left[\vec{Q}\times\left(\frac{\vec{O}\times\vec{C}}{|\vec{O}\times\vec{C}|}\right)\right] = \frac{|V_T|}{r}[\vec{Q}\times(\vec{O}\times\vec{C})]$$

$$= \frac{|V_T|}{r}\begin{pmatrix}\sin\beta\sin\alpha\sin\beta_C\sin\alpha_C + \sin\beta\cos\alpha\sin\beta_C\cos\alpha_C \\ -\cos\beta\sin\beta_C\cos\alpha_C \\ -\cos\beta\sin\beta_C\sin\alpha_C\end{pmatrix} \; . \qquad \text{(A4)}$$

The combined vector, as before, is simply the sum of the translational and rotational velocities:

$$\vec{\omega}(\alpha,\beta) = \vec{\omega}_{\mathrm{T}}(\alpha,\beta)+\vec{\omega}_{\mathrm{R}}(\alpha,\beta)$$

$$= \frac{|V_{\mathrm{T}}|}{S}\begin{pmatrix}-\sin^2\beta \\ \cos\beta\sin\beta\cos\alpha \\ \cos\beta\sin\beta\sin\alpha\end{pmatrix} + \frac{|V_{\mathrm{T}}|}{r}\begin{pmatrix}\sin\beta\sin\alpha\sin\beta_{\mathrm{C}}\sin\alpha_{\mathrm{C}}+\sin\beta\cos\alpha\sin\beta_{\mathrm{C}}\cos\alpha_{\mathrm{C}} \\ -\cos\beta\sin\beta_{\mathrm{C}}\cos\alpha_{\mathrm{C}} \\ -\cos\beta\sin\beta_{\mathrm{C}}\sin\alpha_{\mathrm{C}}\end{pmatrix}. \tag{A5}$$

Observe that if the point Q coincides with C, the point about which the observer is moving, then $S = r$, $\alpha = \alpha_{\mathrm{C}}$, and $\beta = \beta_{\mathrm{C}}$. It follows that the two flow components in equation (A5) will be equal but opposite in sign, giving a resultant optical velocity of 0, as it should be.

This completes the specification of the velocity field relative to the objective sphere for all observer motions. What remains is to specify the optical flow relative to the conceptual retina. Because this requires a knowledge of how the eye actually performs a rotation in kinematic terms, we shall indicate only in a general way what is involved.

Both the position vector (α,β) and the velocity vector (ω) have to be transformed several times to give the optical flow relative to a foveally-based coordinate system on the retina. The first transformation consists of going from object space (the objective sphere) to image space (the conceptual retina). The position vector is modified simply by projecting through the center of the sphere. Thus,

$$\alpha' = \pi+\alpha\,, \qquad \beta' = \pi-\beta\,.$$

The three Cartesian components of the velocity vector are merely inverted in sign. The next transformation accounts for the instantaneous rotation of the globe. As with rotations on the objective sphere, the rotational vector for each position of the globe is added vectorially to the flow vector at that point; this rotational vector is determined by the magnitude and direction of the rotation of the globe relative to the orbit. Finally, the position and retinal flow vectors for each globe position are transformed to a foveally-based coordinate system using the instantaneous position of the retina relative to the orbit.