

Barton GE, Berwick RC and Ristad ES (1988) *Computational Complexity and Natural Language*. Cambridge, MA: MIT Press.

Beesley KR (1996) Arabic finite-state morphological analysis and generation. In: *Proceedings of the 16th International Conference on Computational Linguistics*, Copenhagen, Denmark.

Gildea D and Jurafsky D (1996) Learning bias and phonological-rule induction. *Computational Linguistics* 22(4): 497–530.

Kaplan RM and Kay M (1994) Regular models of phonological rule systems. *Computational Linguistics* 20(3): 331–378.

Karttunen L (1994) Constructing lexical transducers. In: *Proceedings of the 15th International Conference on Computational Linguistics*, Kyoto, Japan.

Kiraz GA (2000) Multi-tiered nonlinear morphology using multi-tape finite automata: a case study on Syriac and Arabic. *Computational Linguistics* 26(1): 77–105.

Koskenniemi K (1983) *Two-level Morphology: A General Computational Model for Word Form Recognition and Production*. Publication no. 11, Department of General Linguistics, University of Helsinki, Finland.

Koskenniemi K and Church K (1988) Complexity, two-level morphology and Finnish. In: *Proceedings of COLING-88*, pp. 335–339. Budapest, Hungary.

Oflazer K, Nirenburg S and McShane M (2001) Bootstrapping morphological analysers by combining human elicitation and machine learning. *Computational Linguistics* 26(1): 59–85.

Sproat R (1992) *Morphology and Computation*. Cambridge, MA: MIT Press.

Theron P and Cloete I (1997) Automatic acquisition of two-level rules. In: *Proceedings of the 5th Conference on Applied Natural Language Processing*, Washington, DC, USA.

### Further Reading

Oflazer K (1999) Morphological analysis. In: van Halteren H (ed.) *Syntactic Wordclass Tagging*. Dordrecht, Netherlands: Kluwer Academic Publishers.

Ritchie GD, Russell GJ, Black AW and Pulman SG (1992) *Computational Morphology*. Cambridge, MA: MIT Press.

Grzywacz, N.M., and D.K. Merwine (2003) Neural Basis of Motion Perception, in the Encyclopedia of Cognitive Science, Vol. 3, 86-98. Macmillan Press, Cambridge, United Kingdom.

# Motion Perception, Neural Basis of

Introductory article

Norberto M Grzywacz, University of Southern California, Los Angeles, California, USA  
David K Merwine, University of Southern California, Los Angeles, California, USA

## CONTENTS

*Introduction*

*Directional selectivity in the retina*

*Directional selectivity in the cortex*

*Models and mechanisms of directional selectivity*

*Selectivity to speed*

*Selectivity to complex motions*

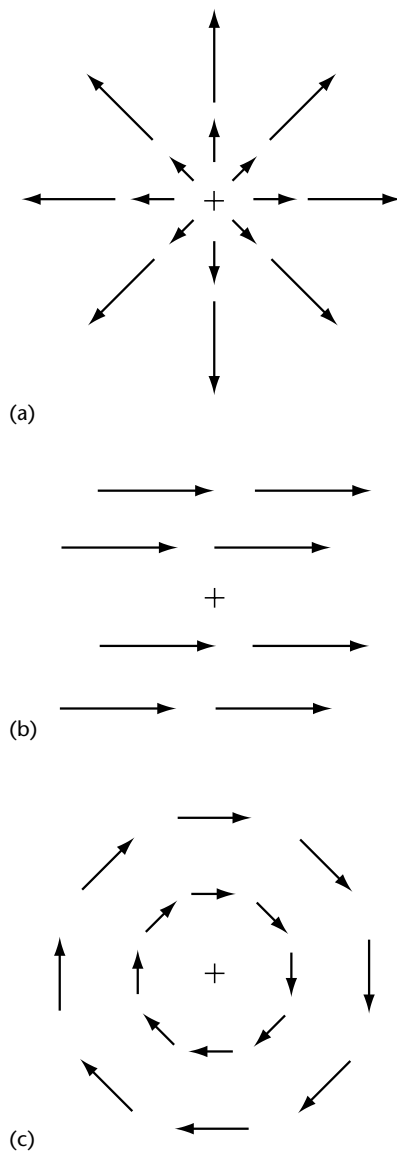
*Animals perceive motion by extracting velocity information from their visual inputs. Differing aspects of the information are computed in a hierarchical series of sequential stages from the retina through the temporal cortex.*

## INTRODUCTION

For most animals, the ability to perceive motion is vitally important. For example, a predator must have the ability to track moving prey in order to hunt and survive, whereas prey must detect the smallest movement of a potential predator. Another example is that of locating and

communicating with a potential mate, which often involves motion detection. Self-navigation is also critically dependent on one's perception of motion relative to surrounding objects. It is therefore not surprising that the ability to perceive motion appears throughout evolution. Animals as simple as the fly contain motion detectors within their visual systems. Most vertebrates are able not only to detect motion, but also to extract its parameters. This article will focus on vertebrate vision. (*See Motion Perception, Psychology of*)

According to Newton, in order to know the motion of a point (or particle), one needs to measure three things, namely direction, speed and



**Figure 1.** Examples of optic flow. The arrows indicate the direction and speed of motion at each point in space. (a) Expansion. (b) Translation. (c) Rotation.

acceleration. The first computation related to motion direction occurs within the eye, and this information is sent to both cortical and subcortical areas of the brain. Within the visual cortex, motion direction is independently determined again from non-directional inputs, and is subsequently refined through a succession of cortical stages. However, retinal and most cortical neurons are not speed selective. Some individual neurons in higher cortical areas, such as the middle temporal cortical area (MT), appear to be speed selective. However, it is generally believed that speed is determined by the combined responses of several neurons. Biological systems are extremely good at determining

the direction of a motion, and are quite good at determining its speed. However, they are poor at determining acceleration, an issue that we shall not explore further here.

One of the most fundamental sources of visual motion is an animal's own movement, namely egomotion. The overall motion of the visual field that results from egomotion is called *optic flow*. For example, if one moves forward in a straight line, the visual flow will spread out from the center-of-heading (Figure 1a), a type of optic flow that is termed *expansion*. In contrast, self-movement backwards results in *contraction* of the optic flow. Sideways movements yield *translation* (Figure 1b), and tilting the head sideways yields *rotation* (Figure 1c). Figures 1a and c illustrate how despite being globally coherent, local-motion directions can be very different and even opposite. Such visual inputs are termed *complex motions*. Many neurons at the higher stages of the motion-processing pathway selectively respond to these types of complex motions.

Motion perception is computed hierarchically – that is, in successive stages within the brain. This article will follow the hierarchy. The computation begins in the retina with a local computation related to the direction of object movement. In the first visual cortical center, namely the primary visual cortex (V1), motion direction is computed again, independently and somewhat differently. From there, a subset of cells sends information to the MT, and from the latter to the middle superior temporal cortex (MST), and so on. The response properties of the neurons along this path are progressively refined. Thus cells in the higher cortical areas respond specifically to complex aspects of motion, such as expansion of the optic flow or object rotation in three-dimensional space. (See **Modularity in Neural Systems and Localization of Function**)

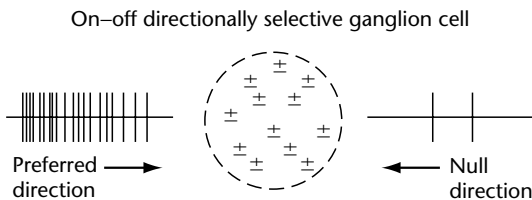
## DIRECTIONAL SELECTIVITY IN THE RETINA

### Basic Description

The electrophysiological study of retinal information processing began with H. Keffer Hartline, who later won the Nobel Prize for his research. He discovered that ganglion cells (the retinal output) would alter their firing rates in response to local changes in brightness. The area in which these luminance changes could influence the firing of a cell was termed the receptive field (RF). Horace B. Barlow subsequently made the first reports of

motion-selective ganglion cell responses in frog retinas in 1953. He found that there were many cells in the frog retina that would fire continuously so long as a visual object moved within the cell's RF. (Interestingly, when the target neurons of a motion-detecting neuron were electrically stimulated, the frog would snap at the corresponding point in space. The motion-detecting neurons were therefore termed 'bug detectors'. This was one of the first demonstrations of a link between neural activity and behavior.) Later, together with Richard M. Hill and William R. Levick, Barlow recorded the first directionally selective (DS) motion responses in a mammalian retina. (*See Single Neuron Recording; Receptive Fields*)

Barlow and his colleagues found two types of DS cells in the rabbit retina. The commonest one is the On–Off DS cell. An On–Off DS cell will respond weakly to a small spot that is flashed on or off anywhere within its RF. However, it will fire strongly if a spot (light or dark) is moved through its RF in an appropriate, 'preferred' direction (Figure 2). Motions in the opposite or 'null' direction yield essentially no response, while orthogonal motions yield intermediate responses. The On–Off DS cells are divided according to their preferred directions into up, down, left and right subgroups, each of which independently tiles the retinal surface. The other DS cell type is the on DS cell. These cells will only respond to bright objects, and they prefer larger objects and slower speeds than do the On–Off DS cells. There are three subgroups of on DS cells, whose preferred directions of motion align with the semicircular canals of the vestibular (balance) system.



**Figure 2.** On–Off DS ganglion cell responses. The rabbit retinal on–off DS cell will respond to both the onset and offset of a small spot flashed within its RF (the area enclosed by the broken circle), as indicated by the  $\pm$  symbol. This cell will respond vigorously to a spot, slit or edge moved in its preferred direction (to the right in this example), but will respond poorly to motions in the opposite or null direction. Upward or downward motions elicit intermediate responses. In this and subsequent figures, the vertical lines denote spikes.

## Species Comparison

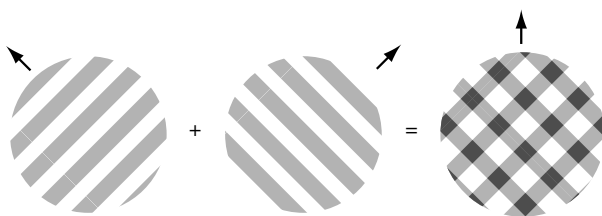
The retinas of mammals, birds, reptiles and amphibians have all been shown to contain DS cells (fish are the only major vertebrate class in which DS cells have rarely been reported). However, as with the two rabbit DS cell types described briefly above, there are many differences in DS cell properties between species. Preferences for light versus dark objects, slow versus fast speeds, sensitivity to object shape and level of dendritic ramification all vary according to the species. Even the percentage of DS cells in the retinal population varies widely. Directionally selective cells are reported to represent between 2% (in cats) and 40% (in turtles) of the retinal ganglion cell population. These percentage differences largely reflect the differences in neocortex available to these species for processing visual information. Thus species with little or no neocortex devote many retinal neurons to computing directional selectivity, while those with large neocortices do not. Nevertheless, despite differences in detail, there is a highly conserved plan among vertebrate species with regard to the use of DS cell output. For all species this information is sent to the accessory optic and pretectal brainstem nuclei. These structures participate in the vestibulo-ocular and optokinetic systems which stabilize the eyes during rotatory head movements, or during rapid global image movements on the retina. Thus these systems are crucial for determining whether image motions on the retina are due to eye, head or body movements or reflect the movement of external objects.

## Detailed Response Properties of On–Off DS Cells

Many studies have been undertaken to elucidate further the properties of the rabbit retinal On–Off DS cell, the aim being to understand the mechanisms responsible for its behavior. For example, it has long been known that an object does not have to pass through the entire RF of an On–Off DS cell in order to generate a DS response (the opposite applies to simple cortical DS cells). Anatomically, the dendritic trees of the On–Off DS cell contain many 'loops', which are consistently around 40–50  $\mu\text{m}$  in diameter. Therefore it was proposed that these cells contained multiple 40–50  $\mu\text{m}$  subunits, each of which performed the DS computation independently. This proposal was consistent with earlier measurements by Barlow and Levick, who attempted to determine the shortest motions that

produce DS responses. However, it was recently demonstrated that these cells can discriminate direction of motion for movements as small as  $26''$  of visual angle (c.  $1\ \mu\text{m}$  on the retinal surface!). This result severely limits the possible mechanisms for computing directional selectivity, implying that they act very locally, they probably involve only a few synapses, and they exist essentially everywhere within the cell's dendritic tree (or those of its inputs). At the time of writing, the function of the dendritic loops remains unclear. (See **Neurons, Computation in**)

Another interesting feature of these cells' responses is that their directional selectivity is invariant to other aspects of a stimulus. For example, consider speed. It is true that an On-Off DS cell will respond best to motions within a narrow speed range. Nonetheless, this cell will respond significantly better to a preferred-direction motion than to a null-direction motion, regardless of the speed. The same is true for temporal frequency. Responses to moving sine- and square-wave gratings may be better for certain temporal frequencies than for others, but the cell always responds better to a preferred-direction stimulus, regardless of the temporal frequency. Finally, imagine that one presents an On-Off DS cell with two simultaneous, non-parallel, drifting sine- or square-wave gratings (so that their overlap forms a 'plaid', as shown in Figure 3). In this case, the cell will respond best when the 'plaid' motion is in the preferred direction. In other words, retinal DS responses are independent of the orientation of the gratings. (This is similar to pattern cells higher in the cortical motion pathway; see below.) In essence, if the cell can 'see' the motion stimulus, it will respond to it in a directionally selective manner. Thus it seems that this cell sacrifices information regarding the exact location, speed, size and shape of stimuli, so that it may robustly indicate their motion direction. The 'lost' information is encoded along the cortical motion pathway.



**Figure 3.** Plaid motion. Two moving gratings are superimposed to form a moving plaid.

## DIRECTIONAL SELECTIVITY IN THE CORTEX

### The Motion Hierarchy

Inputs from the retina pass through the thalamus on their way to the primary visual cortex, also known in primates as V1 (Visual Area 1). From V1, visual information is split along two primary pathways. One path extends into the temporal cortex and contributes to the recognition of objects. The other path, which we shall discuss here, is extended into the parietal cortex and is used to analyze visual motion. This pathway is hierarchical – that is, the information is processed in a series of stages. At each stage, the information is progressively refined. Thus cells at the beginning of the hierarchy respond only when presented with an appropriately oriented, moving edge in a well-defined spatial location. In contrast, cells at higher stages in the hierarchy respond to complex motions, regardless of spatial location. However, it should be noted that the computation of motion is not simply sequential. There is crosstalk with other cortical as well as subcortical areas, and there are back-projections to earlier stages along the hierarchy. Recent experiments suggest that these connections may be crucial for attention, visual awareness and image segmentation (the parsing of the image into regions of relatively homogeneous properties). (See **Parietal Cortex**)

The parietal-motion pathway begins with signals from directionally selective V1 cells. These cells project to the middle temporal cortex (MT or V5) and to V2, whose neurons also project to the MT. From the MT, the motion pathway proceeds to the middle superior temporal cortex (MST). In turn, MST cells project to the lateral and ventral intraparietal areas (LIP and VIP), as well as to V7a. These latter stages in the pathway also carry information for motor planning and control, which are among the main goals of the motion-processing system.

### V1 Cells

The first DS cells along the cortical path appear in the input layers of V1. Nearly all of the so-called 'simple' cells there show some directional selectivity. Indeed, nearly all neurons in V1 are directionally selective to some degree. However, this selectivity is weaker than that in the retina. Moreover, the simple-cell selectivity is dependent on the size, location, spatial frequency and orientation of

the visual stimulus. A simple cell contains separate on and off subregions and responds weakly to a spot flashed in these locations. The best response of this cell occurs when an oriented edge passes entirely through the cell's RF in its preferred direction. As with retinal DS cells, the preferred direction of motion cannot be explained by any spatial organization of the cell's dendritic tree. However, it is often possible to predict the preferred direction of motion for a simple cell by examining its response to flashed spots. Thus directional selectivity in these cells has been explained using a simple additive scheme, which we shall discuss in the section on models and mechanisms of directional selectivity below.

The simple cells project to the upper and lower layers of V1 where 'complex' DS cells are found. These cells do not respond to flashed spots, nor do they have separate on and off subregions. Like simple cells, complex cells prefer oriented edges moving in a particular direction. However, it is no longer necessary for an object to pass through the complex cells' entire RF in order to generate a DS response. It is still unclear to what extent (if at all) the directional selectivity of simple cells contributes to that of complex cells.

## MT Cells

A subset of DS V1 cells projects to the MT. Cells in the MT have large RFs, often 10 times larger than those of cells in V1. MT cells will respond to an object moving in their preferred direction anywhere within their RF. In addition, J. Anthony Movshon and colleagues found that approximately one-third of the cells in the MT (called pattern cells) could detect the motion of a plaid in a DS manner. That is, suppose that the cells are presented with two overlapping sine waves moving in different directions (Figure 3). In this case, these cells will respond best when the composite motion of the sine waves, not their individual motions, is in the preferred direction. This behavior is illustrated in Figure 4. In contrast, the component MT cells (the other two-thirds) behave similarly to cells earlier in the cortical hierarchy – that is, they respond whenever either of the two gratings is moving in the preferred direction of the cell. Not surprisingly, human subjects see the coherent plaid motion under many stimulus conditions (i.e. subjects often see the motion 'reported' by the pattern cells). However, under some conditions (e.g. when the sine waves have very different contrasts and spatial frequencies), subjects see the sine waves sliding past each other. This percept is termed transparency.

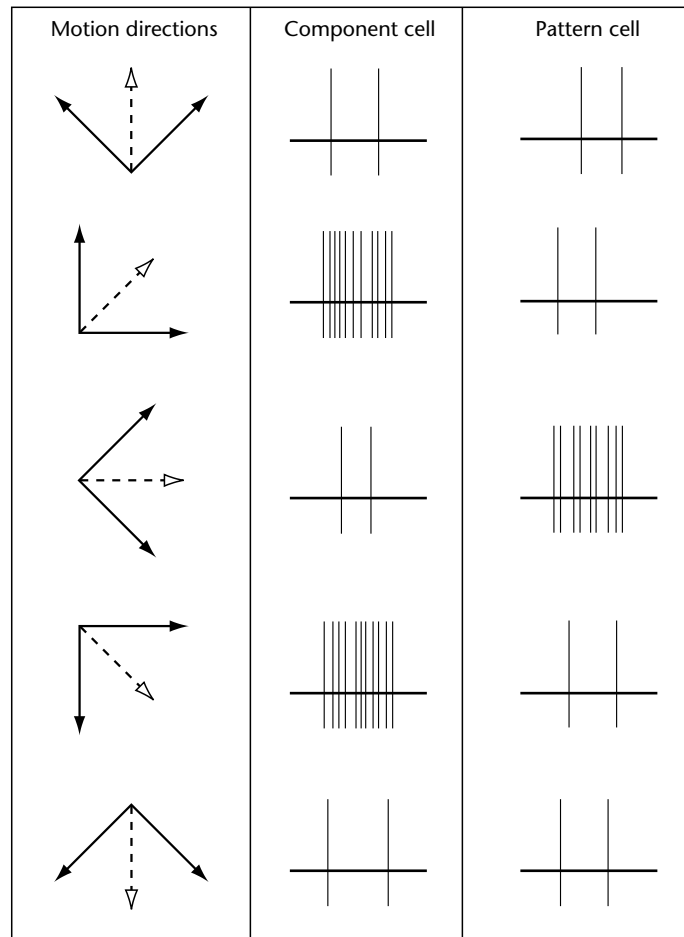
## MST Cells

Cells in the MT primarily send their outputs to the middle superior temporal cortex (MST). Here DS information from several cells converges to create neurons with very large RFs. Neurons in the MST are the first motion neurons in which RFs are bilateral – that is, they extend across the visual midline. Some neurons in the dorsal region of the MST have RFs that cover most of the visual field. These cells combine DS inputs in such a way that they are sensitive to complex visual motions, such as contraction, expansion or rotation of the optic flow. A subregion of the MST is also devoted to encoding translational motions. Interestingly, the cells that respond to translational motions are active so long as the stimulus is present. Cells that are attuned to rotational or expansive motions, although they show a strong sustained component to their response, also display a strong transient response to the onset of a complex motion. This signal may prove crucial for navigating through the environment while simultaneously moving one's head and eyes. We shall discuss these responses in more detail in a later section.

Before we leave this section, we would like to mention an area located lateral and anterior to human MT/MST. (This newly discovered area and the MT/MST complex are both in the superior-temporal sulcus.) This area is interesting because it extends the capabilities of the MST to the detection of biological motions, such as 'point-light walkers'. Point-light walkers are images produced by placing lights on the joints of a moving person. Subjects readily recognize these displays, and can even report the gender of the 'walker'. The newly discovered area is specifically activated by such displays.

## Correlation between Cell Responses and Perception

William T. Newsome and colleagues have performed a fascinating series of experiments with MT neurons. The aim was to investigate whether motion perception is genuinely determined by their activity. These scientists trained monkeys to report their perception of the direction of a group of moving dots. Some of these dots were correlated (i.e. moving in the same direction), while others were not. By varying the percentage of correlated dots, Newsome and colleagues discovered that the monkeys' discrimination mirrored directly the activity of the cells in the MT. Furthermore, as the responses varied because of noise, the



**Figure 4.** Responses of component and pattern MT cells. The left-hand column shows motion stimuli, while the center and right-hand columns show the responses of component and pattern MT cells, respectively. For the motion stimuli, the solid arrows indicate the motion directions of two identical sine-wave gratings. The broken, open arrows show the motion of the plaid formed by the combination of the gratings (Figure 3). A component cell will respond whenever either grating is moving in the preferred direction (to the right for both cells in this figure), as shown in the second and fourth rows of the second column. This cell does not respond when the plaid motion is in the preferred direction (third row of second column). A pattern cell, on the other hand, will not respond to the individual gratings, but will respond to the plaid motion.

discrimination also varied in a predictable manner. These researchers further showed that they could control the monkeys' reported perception by injecting small amounts of current into the MT during the experimental trials. Thus they provided strong evidence that the firing rate of these neurons directly encoded the animals' perception.

## MODELS AND MECHANISMS OF DIRECTIONAL SELECTIVITY

### Models of Directional Selectivity

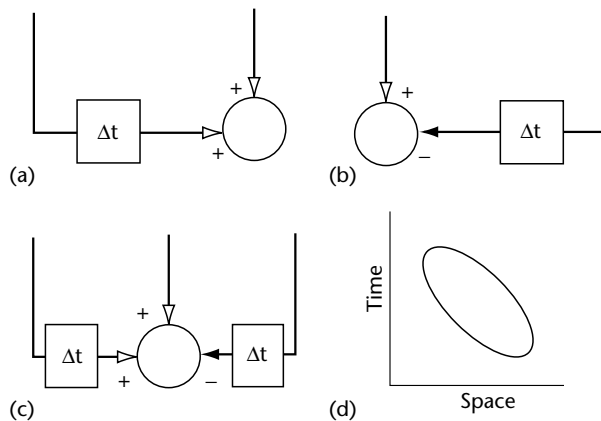
During their studies of retinal directional selectivity in insects, Werner T. Reichardt, Tomaso Poggio and their colleagues described the theoretical

requirements for any model of directional selectivity. The first requirement is a spatial asymmetry. That is, if a neuron responds better to motions coming from the left than to motions coming from the right, then there must be some difference in the neural inputs from the left and right sides of the cell's RF. Essentially, all DS models propose that this asymmetry is temporal. In other words, some difference in time course exists between the left- and right-side inputs. However, this need not be the case. For example, the left-side input could 'gate' the right-side input. In this case, the cell will only fire if the motion comes from the left and opens the gate before the right-side input is activated.

The second requirement for producing directional selectivity is a nonlinear mechanism (i.e. a

mechanism that does not simply additively sum its inputs to determine its output). A spatial (plus temporal) asymmetry as described above can generate a directional difference in the response (i.e. differing response waveforms depending on the direction of stimulus motion). However, this alone is not sufficient to produce two different single-number responses for preferred- and null-direction motions. The work by Poggio and Reichardt proves that, without a nonlinear mechanism, the numbers obtained by integrating (summing over time) the waveforms of the responses must be equal for all directions of motion. Some nonlinearity, perhaps as simple as a threshold, must be present. (A threshold nonlinearity operates by only allowing responses when the input exceeds some minimum value.)

Figure 5a illustrates the simplest model proposed for insect retinal directional selectivity,



**Figure 5.** Models of theoretical DS mechanisms. In the retinal models (a, b and c), the lines with arrows represent inputs to nonlinear interaction sites (denoted by circles). Boxed  $\Delta t$  symbols denote slow or delayed input lines. Movement proceeding from the slow line to the fast line can create signals that arrive at the interaction site simultaneously. In this figure, these signals are such that all of the illustrated models prefer rightward motions. (a) In the Reichardt model, the proposed interaction is facilitatory (multiplication). (b) In the Barlow and Levick model, the proposed interaction is inhibitory (now believed to be division-like). (c) In the two asymmetrical pathways model, both facilitation and inhibition operate, allowing for robust DS to a variety of motion stimuli. (d) Spatio-temporal inseparability. The horizontal axis indicates the spatial location of a visual stimulus, while the vertical axis shows the time of peak response. For a cell with spatio-temporally inseparable responses, the tilted oval bounds the response region. Thus there is an orderly progression between space and time, such that only for motion in a particular direction will the responses from several points in space coincide in time.

known as the Reichardt model. For clarity, the model uses inputs from only two locations, although more complex versions are possible. (The mechanism is presumably replicated many times within the RF of the DS cell.) The proposed spatial asymmetry in this model is temporal. Inputs from the side first encountered by an object moving in the preferred direction (left in the figure) propagate to the interaction site slowly compared with those from the null side. The proposed nonlinearity for their interaction is multiplication. Therefore if an object moves in the preferred direction at the right speed, the slowness of the left-side path is compensated for by the earlier arrival time of the stimulus. This allows both inputs to arrive at the interaction site at the same time, yielding a positive multiplication. For null-direction motions the inputs will arrive separately, and the result of the multiplication will be zero. Interestingly, many psychophysical models of human motion perception follow variants of this multiplicative Reichardt model. (See **Computational Neuroscience: From Biology to Cognition**)

## Retinal Mechanisms for Directional Selectivity

As part of their study of rabbit DS cells, Barlow and Levick performed a series of two-slit apparent-motion experiments. These scientists drew attention to the responses to two adjacent slits that were flashed as if an object was moving in the null direction. The response generated by the DS cell under study was far less than the sum of the responses for each slit flashed alone. Barlow and Levick concluded from this that directional selectivity in these cells was produced by a nonlinear, inhibitory mechanism that ‘vetoes’ responses to null sequences. As shown in Figure 5b, their proposed spatial asymmetry has two components. The central component is excitatory and propagates quickly. The second, inhibitory component is offset to the null side and propagates slowly. Thus an asymmetry exists in both the sign and the time course of the two spatially separated components. Therefore motions in the preferred direction will yield responses, while null-direction motion responses are vetoed. Studies performed more recently by Franklin R. Amthor and colleagues have shown that these cells cannot perform a perfect veto, and that a better description of the inhibitory interaction is a division-like nonlinearity.

In addition, the response of a DS cell to a preferred-direction apparent motion is greater than

the sum of the responses to each slit alone. This is called preferred-direction facilitation. Under some circumstances, preferred-direction facilitation can be as strong as null-direction inhibition. The excitation responsible for facilitation appears to originate from starburst amacrine cells which release the neurotransmitter acetylcholine. It has been shown that there is a spatial asymmetry in the input-output relationship of the dendritic trees of these cells. Their dendrites receive excitatory inputs along their entire length, but they release excitatory neurotransmitter and may receive inhibitory inputs only at their tips. (These neurons do not have axons!) Several lines of evidence indicate that the tip inhibition acts in a division-like manner. Therefore each dendrite contains a spatial asymmetry and a nonlinearity, and can act as an autonomous DS unit. For this reason it has been proposed that DS signals appear presynaptically (before the DS ganglion cell itself), and flow from the dendrites of the starburst amacrine cell to the DS cell. A DS cell could therefore generate its responses by preferentially sampling from starburst-cell dendrites with a common preferred direction.

However, there is accumulating evidence that both pre- and postsynaptic asymmetries may be involved in producing retinal directional selectivity. Complete blockade of starburst-cell outputs with acetylcholine antagonists does not entirely eliminate this selectivity to moving bars. The residual direction selectivity appears to be postsynaptic. In turn, blocking the inhibitory input to a DS cell (with antagonists of the neurotransmitter GABA) does not always eliminate retinal directional selectivity. Moreover, such blockade occasionally even reverses the cell's preferred and null directions. Computer simulations of the starburst dendritic directional selectivity can account for these reversals because of synaptic saturation. Thus asymmetrical postsynaptic inhibition and asymmetrical presynaptic facilitation may act cooperatively (Figure 5c) to produce robust directional selectivity in On-Off DS ganglion cells.

### Cortical Mechanisms for Directional Selectivity

There are two classes of models that have been proposed to account for cortical directional selectivity. One class, advanced by George Sperling and colleagues, is based on human psychophysics, and is similar to the Reichardt model shown in Figure 5a. Another class consists of the motion-energy models advanced by Edward H. Adelson and James R. Bergen. As with all DS models,

motion-energy models require a spatial asymmetry. The proposed spatial asymmetry is that successive adjacent locations in the cell's RF will respond with gradually decreasing sluggishness. The cell's RF profile is therefore tilted in space and time, as shown in Figure 5d by the slanted oval. This property is known as space-time inseparability. As can be seen in the figure, for this type of space-time arrangement only one direction of object motion (to the right) can result in the DS cell's inputs all arriving simultaneously. As before, simple linear summation will result in differential response waveforms for preferred- and null-direction motions, and some nonlinearity must exist to convert this directional difference into directional selectivity. The most commonly proposed nonlinearity is squaring, which is described as extracting the motion energy from the directional difference. (See **Psychophysics**)

Robert M. Shapley and colleagues obtained physiological evidence for space-time inseparability in both the on and off subregions of the inputs to the simple DS cells found in V1. However, the correlation between the simple cell space-time profile and direction selectivity varies widely in V1. Cells in some layers of V1 show a very high correlation, while those in other layers show a very low correlation, despite equivalent directional tuning. Thus space-time structure alone cannot completely account for simple DS cell responses. In addition, these models of DS simple cells generally overestimate non-preferred responses and sometimes underestimate preferred responses. Moreover, these models do not predict onset transients, which are commonly observed. Therefore both inhibitory and excitatory feedback interactions between cortical cells have been proposed to account for these discrepancies. The exact mechanisms for producing directional selectivity in these cells are still the subject of debate.

Because the complex DS cells of V1 do not respond to flashed spots, they cannot show space-time-oriented RFs. However, it has been demonstrated that the interactions between two sequentially stimulated locations in a complex cell's receptive field are space-time inseparable. This behavior is termed second-order space-time orientation. Dynamic nonlinearities have been proposed to account for the directional selectivity of complex DS cells. These nonlinearities would facilitate or inhibit, respectively, the responses to preferred- or null-direction motions. Similarly, space-time-separable simple cells have also been shown to display some second-order space-time structure. Because there is ample evidence for



interactions between complex and simple cells, it is possible that second-order space–time inseparability in simple cells arises from complex-cell inputs.

The directional selectivity in MT neurons is more sophisticated than that in V1, especially in pattern cells (Figure 4). Many models have been advanced to account for the orientation independence of pattern cells. These models often begin by pointing out the aperture problem of V1 cells – that is, the difficulty in determining the true direction of motion through small apertures (e.g. small RFs). This is because they only reveal small, straight portions of contours, and it is impossible to tell the direction of motions parallel to straight edges. Models of pattern cells combine the responses of V1 (or component) cells with many preferred directions to disambiguate the direction of motion of multi-orientation patterns.

## SELECTIVITY TO SPEED

The visual system measures not only the directional component of the velocity vector but also its magnitude – the speed. Measuring speed is important for vision. For example, humans appear to use both direction and speed signals to obtain a precision of 1–2° in the estimation of heading direction. Furthermore, Richard Andersen and colleagues have shown that both humans and primates benefit from speed information when estimating the three-dimensional structure of objects from motion cues. Psychophysical findings also indicate that one can segment images based solely on gradients of speed signals. Moreover, smooth pursuit and saccades are made precise because speed signals are used in their computations. Finally, the visual system must somehow use speed signals to achieve effective deblurring. (Motion blurring occurs when one leaves the shutter of a camera open for too long when there is motion in the image.)

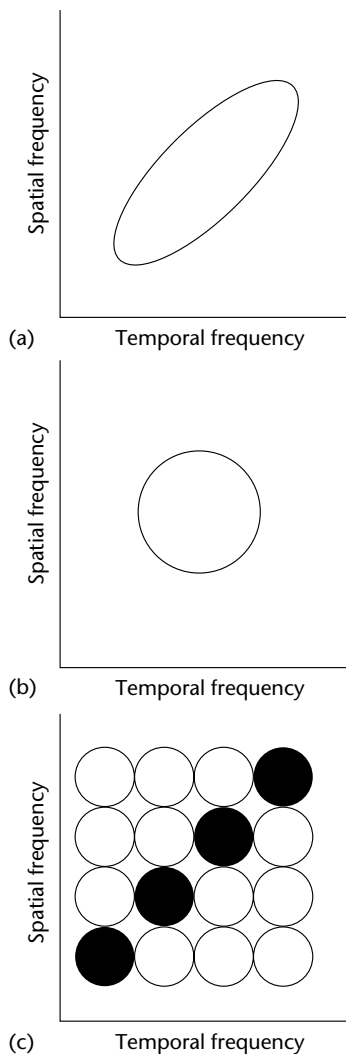
However, measuring speed is more difficult than measuring the direction of motion. It is difficult to measure the former because it is the magnitude of the derivative of position over time, and thus requires precise spatio-temporal information. In contrast, to measure direction one only requires two relatively imprecise position measurements. Psychophysically determined measurements of local and instantaneous speed have been found to be very noisy. The relative errors in the measurement of local speed are in the range 30–100%.

Nevertheless, humans can measure speed precisely if they are provided with a relatively long trajectory of motion. Under these conditions, the errors made when discriminating speed can be as

low as 5%. This high level of precision seems to be achieved by integrating relatively imprecise local-speed signals over time. The 5% precision of velocity determination occurs in many experimental conditions, including motions of dots, edges, sine-wave gratings (of varying spatial and temporal frequencies), plaids and frequency-modulated stimuli. (In the latter, the local-intensity profile is not moving, but the contrast of local portions of the profile increases transiently, and this perturbation propagates with fixed velocity.)

Perhaps the simplest method of measuring local visual speed is by its derivative definition. This involves finding the positions of an image feature in two discrete instances in time and then computing the ratio between the positional distance and the temporal delay. It is essentially the approach proposed by Shimon Ullman in his minimal mapping theory. He proposed that the main problem facing the visual system when measuring motion is to solve the correspondence problem – that is, to find correspondences between image features at one instant in time and (hopefully) the same features at the next instant in time. Ullman suggested that the features correspond to minimize the total distance traveled. After correspondence has been established, distance, delay and thus the velocity of a feature can then be calculated. Serious challenges to Ullman’s emphasis on the correspondence problem, and thus his method for measuring visual speed, have been raised by motion psychophysicists, who pointed out that his theory was not immediately consistent with known neural processes underlying motion perception. Moreover, they argued that the correspondence problem is essentially non-existent when one is dealing with real neural RFs. (See **Computational Neuroscience: From Biology to Cognition**)

When one looks directly at neural responses for clues as to how the brain measures speeds, a puzzling finding arises. Neurons have a sharp optimal speed if they are stimulated with moving edges, especially in the MT. However, cells in V1 and 40% of the cells in the MT do not seem to detect speeds. If these cells were speed tuned, then if one were to raise the temporal frequency of the stimulus, the spatial frequencies that yield the optimal responses should increase in proportion (Figure 6a). Rather, these motion sensitive cells are tuned to spatio-temporal frequencies, regardless of speed (Figure 6b). Nevertheless, the work of John A. Perone and Alexander Thiele revealed that 60% of MT cells show true speed selectivity. So how do 60% of MT neurons compute speed from inputs that are not speed selective?



**Figure 6.** Theories for speed encoding. For (a), (b) and (c), the vertical axis indicates the spatial frequency of a moving sine-wave grating and the horizontal axis indicates the temporal frequency at which the intensity of every point oscillates. (a) Idealized relationship between optimal spatial and temporal frequencies for a cell that can encode speed. For any given speed, as the spatial frequency of the grating increases, the optimal temporal frequency increases in proportion. Therefore speed is indicated by the slant (slope) of the oval response region. (b) Spatio-temporal profile of a non-speed-selective cell on the motion pathway. Most motion-selective cells are band-pass in both the spatial and temporal domains. Non-speed-selective cells do not show the organized relationship (shown in a) that is required for speed selectivity. (c) A population code for speed. Each cell is represented by a circle whose center denotes the optimal spatial and temporal frequencies of the cell. Solid and open circles illustrate responding and non-responding cells during a translation. By looking at a population of cells with band-pass spatio-temporal response properties, a higher-level unit can determine speed regardless of the spatial and temporal frequencies of the stimulus.

## A Population Code for Speed

Several investigators have proposed that speed selectivity arises from a population code. In such a code, a speed-selective cell high in the motion-pathway hierarchy could read the speed of the stimulus from the simultaneous firing of many earlier cells. Eero Simoncelli, David Ascher and colleagues obtained psychophysical data that support the concept of a population code. We shall now explain this idea in more detail.

Population-code models for the measurement of local speed begin by looking at the responses of a population of DS neurons with different receptive-field sizes. It is known that V1 neurons with increasingly larger receptive-field sizes are tuned to increasingly lower spatial frequencies when they are stimulated with sinusoidal gratings. Moreover, different neurons tend to be tuned to different temporal frequencies of the gratings. The population-code models for local speed measurement consider a three-dimensional space defined by these tunings. In this space, two axes are the optimal horizontal and vertical spatial frequencies that drive different DS cells, while the third axis represents their optimal temporal frequencies. A point in this space corresponds to a DS cell. This space is of interest when a visual translation of given velocity covers the receptive fields of these cells. The optimal responses tend to fall on a plane in this space. Figure 6c illustrates the projection of this plane on to the plane formed by the temporal-frequency axis and one of the spatial-frequency axes. To measure local speed in population-code models, one must detect the slant of the projecting plane relative to the temporal-frequency axis. Several schemes for detecting this plane have been proposed in the literature. These schemes depend on the exact spatio-temporal properties of the cells, although Ascher and colleagues described how to design good schemes under broad conditions.

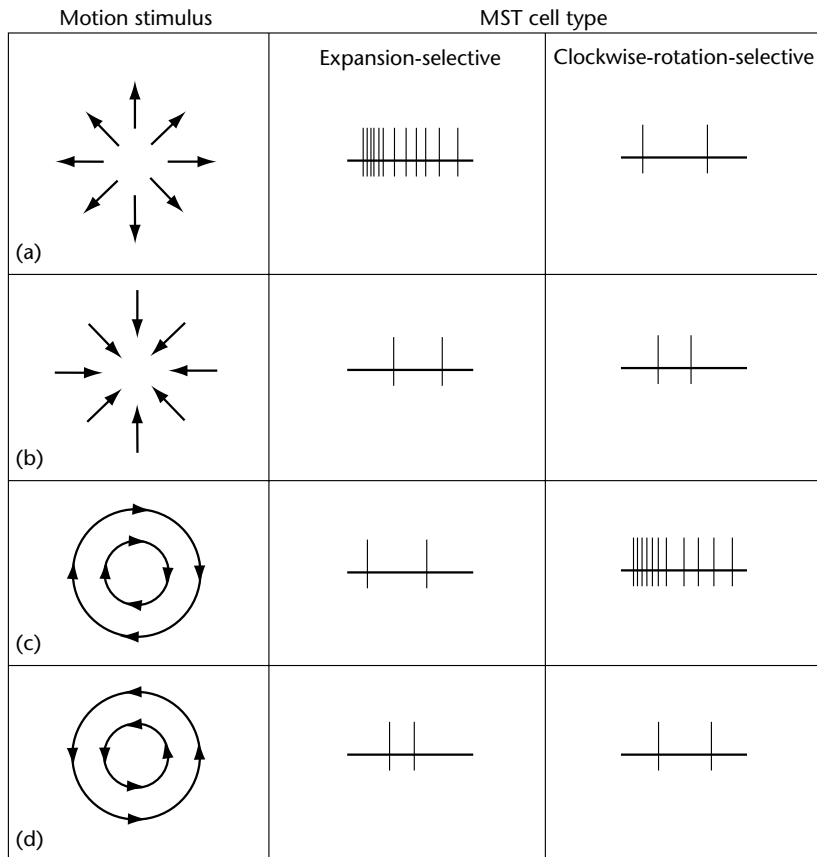
Finally, although we first introduced the concept of population codes in the context of determining speed, such codes must occur generally in the brain. Even if a cell is tuned to a particular property, the response of that cell may be modulated by other properties. This would mean that a cell's firing rate was not unique. Consequently, the brain must look at the firing of several cells in order to disambiguate any particular property. In the motion domain, in addition to speed, such a disambiguation also occurs for direction of motion. For example, one can obtain local direction from the tilt of the plane described above relative to the spatial frequency axes. (*See Decoding Neural Population Activity*)

## SELECTIVITY TO COMPLEX MOTIONS

### Interesting Types of Complex Motion and MST

The two preceding sections focused on the measurement of two local variables of motion, namely direction and speed. However, as was shown in Figure 1, the motion of optic flow is in general globally complex. This complexity arises because, for example, as one moves through the environment, neighbor velocity vectors are statistically dependent. A similar dependence occurs when a rigid or quasi-rigid object moves in front of one's eyes. Jan J. Koenderink showed that if one considers small planar patches in the surface of moving, rigid objects, their general motion can be decomposed in terms of translational, radial (expansion and contraction), rotational and shear motions. These types of motion can be modeled with a few parameters describing the dependence of the local velocity vectors. For instance, one

can describe a rotation by specifying its center and its angular velocity. In turn, one can specify an expansion by specifying the focus and rate of expansion. Thus if the brain could measure the few parameters of translational, radial, rotational and shear motions, then it could estimate such things as direction of egomotion heading. As was explained earlier, the MST is the first cortical area to be selective for these motions. Neurons in the dorsal portion of this area (MSTd) respond selectively to these motions, either alone or in combination (e.g. spiral motions). Figure 7 illustrates how one sees this selectivity in the electrophysiological recordings. An expansion-selective cell (second column) responds strongly to expansion (a) but not to contraction (b) or rotations (c and d). In contrast, a cell that is selective for clockwise motion (second column) responds strongly to this motion (c) but not to counter-clockwise rotation (d) or radial motions (a and b). These cells also do not respond to translation (not shown in the figure).



**Figure 7.** Responses to complex motions by MST cells. The left-hand column shows various complex-motion stimuli, while the center and right-hand columns indicate the responses of expansion-selective and clockwise-rotation-selective cells, respectively. Each cell type responds only to a specific global coherent motion presented within its RF. (a) Expansion. (b) Contraction. (c) Clockwise rotation. (d) Counter-clockwise rotation.

## Roles of MST Neurons

What are the behavioral functions of the MST cells that are selective for complex global motions? Kenneth J. Britten and Richard J. van Wezel showed that microstimulation of neurons in this area influences the heading behavior of primates. Related to this, Andersen and colleagues showed that many MSTd neurons shift their focus-of-expansion tuning curves, compensating for retinal motions during eye movements. (This tuning curve is the cellular activity as a function of the focus location of the stimulus.) Because the focus of expansion indicates the direction of heading, this and the results of Britten and Wezel demonstrate the contribution of these neurons to this function. Another of their roles was demonstrated by lesions of the MST, which impaired the animal's ability to execute a smooth-pursuit eye movement when the target moved towards the lesioned hemisphere. (A similar deficit was seen for optokinetic nystagmus movements.) Furthermore, microstimulation within the MST influenced the velocity of smooth-pursuit eye movements. Thus MST neurons must also contribute to the control of eye movements. A final role that we shall mention here is that MST neurons can code the abstract concept of a motion as well as the actual motion of a visual stimulus. In the experiment demonstrating this role, animals saw some stimuli appear and then move in some trials. In other trials, the stimuli appeared, disappeared, and then reappeared at the same final location as the moving stimuli. This simulated a motion that might have occurred behind an occluder. Many MST neurons coded the direction of the occluded motion, which strongly suggested that the MST has a central role in the perception of motion.

## Mechanisms of MST Receptive Fields

It is thought that the receptive field properties of MST neurons emerge from the combination of properties of earlier neurons. This is very similar to pattern cells in MT being built out of component cells. For example, one could build a clockwise-rotation-selective cell by using DS cells with leftward, upward, rightward and downward preferred directions. All one has to do is to place these cells in the bottom, left, top and right visual fields, respectively. However, one problem with this idea is that these cells show position and scale invariance. In other words, although the magnitude of the response may vary with location, the selectivity for a particular pattern of motion remains the same throughout the receptive field, regardless of size.

Thus building MSTd RFs is not simple, and may involve dendritic subunits such as those described earlier for the retina. A further complication is that the complex selectivities in MST are also independent of the cues that convey the motion. For instance, effective expansion stimuli could be generated by illusory contours. Consequently, one must explain how any motion that appears perceptually as an expansion will activate an expansion neuron, even if the luminance pattern is not expanding.

It also remains unclear whether MST computations include estimations of all of the parameters of the optic-flow components described by Koenderink. Koenderink himself, and some of his colleagues, performed psychophysical experiments which showed that humans do not metrically discriminate all of the parameters of complex motions. For example, according to these experiments, humans may not discriminate between angular velocities. However, many of the experiments used impoverished visual displays. Experiments that have been conducted more recently with richer displays indicated that humans could measure quantitatively certain complex-motion parameters, including angular velocity.

## Further Reading

- Andersen RA (1997) Neural mechanisms of visual motion perception in primates. *Neuron* **18**: 865–872.
- Borg-Graham LJ and Grzywacz NM (1991) A model of the direction selectivity circuit in retina: transformations by neurons singly and in concert. In: McKenna T, Davis J and Zornetzer SF (eds) *Single Neuron Computation*, pp. 347–375. Orlando, FL: Academic Press.
- Croner LJ and Albright TD (1999) Seeing the big picture: integration of image cues in the primate visual system. *Neuron* **24**: 777–789.
- Grzywacz NM and Yuille AL (1991) Theories for the visual perception of local velocity and coherent motion. In: Landy MS and Movshon JA (eds) *Computational Models of Visual Processing*, pp. 231–252. Cambridge, MA: MIT Press.
- Grzywacz NM and Merwine DK (2002) Directional selectivity. In: Arbib M (ed.) *The Handbook of Brain Theory and Neural Networks*. Cambridge, MA: MIT Press.
- Miles FA and Wallman J (1993) *Visual Motion and its Role in the Stabilization of Gaze*. Amsterdam: Elsevier Science.
- Smith AT and Snowden RJ (1994) *Visual Detection of Motion*. London, UK: Academic Press.
- Watanabe T (1998) *High-Level Motion Processing: Computational, Neurobiological and Psychophysical Perspectives*. Cambridge, MA: MIT Press.
- Wurtz RH and Kandel ER (2000) Perception of motion, depth and form. In: Kandel ER, Schwartz JH and Jessell

TM (eds) *Principles of Neural Science*, pp. 548–571. New York: McGraw-Hill.

Zanker JM and Zeil J (2001) *Motion Vision: Computational, Neural and Ecological Constraints*. Berlin, Germany: Springer Verlag.

Zeki S (1990) The motion pathways of the visual cortex. In: Blakemore C (ed.) *Vision: Coding and Efficiency*, pp. 321–345. Cambridge, UK: Cambridge University Press.

# Motion Perception, Psychology of

Introductory article

George W Mather, University of Sussex, Brighton, UK

## CONTENTS

*Introduction*

*Uses of motion information*

*Motion cues*

*The motion after-effect*

*Apparent motion*

*Induced motion*

*Motion perception is important for figure–ground segregation, three-dimensional vision, and visual guidance of action. Specialized brain cells detect image motion. Adaptation in these cells leads to illusory motion, such as the motion after-effect.*

## INTRODUCTION

An essential attribute that distinguishes all animals from plants is their capacity for voluntary movement. Animals move to find mates, shelter, and food, and to avoid being eaten. But the ability to move brings with it the need to sense movement, whether to navigate through the world, or to detect the movement of other mobile animals such as approaching predators. For sighted animals, this means sensing movement in the visual image that is projected into the eye. The image is formed on a sheet of light-sensitive cells that line the inside of the eye – the retina. Specialized neural processes are required to detect the presence of movement in the retinal image.

## USES OF MOTION INFORMATION

Surfaces, shapes, and objects in the scene under view create spatial patterns of light and dark in the retinal image. The image is very rarely still, as in a photograph. Instead, it is in a state of continuous change, due to the movement of objects in the scene (e.g. an approaching predator) or to shifts in the position of the observer's eyes, head, or body (e.g. while running away from the predator).

Perception of movement in the image is crucial, because it can be used in a number of ways.

## Figure–Ground Segregation

Shapes and objects that are invisible while static (e.g. camouflaged animals) are revealed as soon as they move relative to the background. Many animals have evolved special ways of moving, in an attempt to defeat figure–ground segregation. For example, prey animals such as lizards and rodents move in short, rapid bursts in between periods of complete stillness, in order to minimize the chances of detection by predators. Predators such as cats tend to move slowly and smoothly to avoid being seen by their prey.

## Extraction of Three-dimensional Structure

When any solid object moves, the images of its various parts that are cast on the retina move relative to each other. Relative motion of this kind can be used to extract the three-dimensional structure of the object. For example, in a sideways view of a rotating globe, surface markings near the equator move across the field of view more rapidly than markings near the poles. In addition, markings near the equator follow almost a linear path, whereas those near the poles follow more elliptical paths. This highly structured variation in speed and direction is sufficient for the perception of the shape's three-dimensional structure.