

Chapter prepared for
Stevens' Handbook of Experimental Psychology, 3e
Hal Pashler, Series editor, Steve Yantis, Volume editor
J. Wiley Publishers, New York

PERCEPTION OF VISUAL MOTION

Robert Sekuler, Scott N. J. Watamaniuk and Randolph Blake

Brandeis University, Wright State University and Vanderbilt University

e-mail to:

sekuler@brandeis.edu

watamaniuk@wright.edu

Randolph_blake@vanderbilt.edu

INTRODUCTION AND OVERVIEW

Gordon Lynn Walls, a comparative anatomist, observed: “If asked what aspect of vision means the most to them, a watchmaker may answer ‘acuity,’ a night flier ‘sensitivity’, and an artist ‘color.’ But to animals which invented the vertebrate eye and hold the patents on most features of the human model, the visual registration of motion was of the greatest importance (Walls, 1942).”

The rich and rapidly-expanding scientific literature on visual motion perception suggests that Walls was right: to organisms all up and down the phylogenetic scale, visual motion perception is of unmatched importance. Visual motion serves a wide variety of crucial roles: wayfinding (optic flow), perception of shape from motion, depth segregation, judgments of coincidence (time to collision, time to filling a tea cup), judgments of motion direction and speed and perception of animate, biological activity. Sometimes, the presence of motion can compensate for deficiencies in other forms of visual information, as Figure 1 shows. The three images in that figure are frames from a video showing a person performing a common action. Clearly, no single frame conveys sufficient spatial structure to permit recognition that a person is present, let alone recognition of what the person might be doing. However, the complex patterns of visual motion generated when these frames are displayed as part of a video convey immediately that a person is present, and that the person is in the process of sitting down (Bobick & Davis, 2001).¹

Recent decades have produced major advances in understanding of visual motion perception.² Many such advances have come from complementary approaches to analyze motion: psychophysical, computational and neurophysiological. It is known now that the detection and analysis of motion are achieved by a cascade of neural operations, starting with the registration of local motion signals within restricted regions of the visual field, and continuing with the integration of those local motion signals into more global descriptions of the direction and speed of object motion. Physiological studies of animals -- most notably cats and monkeys -- have revealed some of the neural hardware comprising this hierarchical processing scheme. And recently, exciting new techniques including brain imaging and transcranial magnetic stimulation have been deployed in concert with psychophysics to identify neural concomitants of motion perception in the human visual system.

Our goal in this chapter is to highlight some of these exciting developments. However, limitations on space -- together with the exponential growth of the literature on motion perception -- forced on us hard choices about what to include or omit. Thus, this chapter emphasizes motion in the front-parallel plane, unavoidably de-

¹The video can be downloaded from
<< <http://www.cis.ohio-state.edu/~jwdavis/Archive/blurmotion.mpg>>> or
<<<http://www.cis.ohio-state.edu/~jwdavis/Archive/blurmotion.mov>>>.

² Previous editions of this handbook paid scant notice to the topic of our chapter. In the first edition, Graham (1951) spent just six pages on motion perception, emphasizing research on apparent motion. In the second edition, more than three decades later, the coverage was increased by only a just noticeable difference, to a total of ten pages distributed over two chapters (Hochberg, 1988; Westheimer, 1988).

emphasizing work on motion in depth and “cyclopean” motion perception (Patterson, 1999). In addition, the chapter focuses on motion of objects defined by luminance contrast, with little discussion of important work on the role of chromatic information in motion processing (Dobkins, 2000; Gegenfurtner & Hawken, 1996). The chapter slights numerous interesting and potentially informative illusions of motion (e.g. Hikosaka, Miyauchi, & Shimojo, 1993; Krekelberg & M, 2001; Viviani & Stucchi, 1989). Moreover, our coverage primarily focuses on motion perception in primates, particularly *Homo sapiens*. Consequently, interesting work on motion perception in birds (e.g. Bischof, Reid, Wylie, & Spetch, 1999; Wylie, Bischof, & Frost, 1998), fish (e.g. Albensi & Powell, 1998; Orger, Smear, Anstis, & Baier, 2000) and insects (e.g. Dror, O'Carroll, & Laughlin, 2001; Gabbiani, Mo, & Laurent, 2001) has been left out. Our chapter does include research on “atypical observers,” particularly individuals with diminished motion sensitivity consequent to brain damage.

Stimuli

In introducing the first edition of this handbook, Stanley Smith Stevens observed that “In a sense there is only one problem of psychophysics, the definition of the stimulus. ...[T]he complete definition of the stimulus to a given response involves the specification of all the transformations of the environment, both internal and external, that leave the response invariant. This specification of the conditions of invariance would entail, of course, a complete understanding of the factors that produce and that alter responses (Stevens, 1951, pp. 31-32).” We agree. As this chapter underscores, contemporary research on visual motion perception has advanced in large measure because of researchers’ ability to generate and deploy suitable stimuli, including innovative computer-generated animations, which simulate complex, real-life events.

Commenting on one aspect of this challenge, Graham cautioned “...[W]e must take care that parameters are not confounded, a danger that arises only too readily from the fact that velocity itself involves the variables of distance and time. In any given experiment the variables of time, distance, interval between stimuli, and cycle of repetition of stimuli must be clearly analyzed before we can be confident that unequivocal conclusions maybe drawn (1965, pp. 579-580).”

Researchers have developed many clever ways around the confounds that Stevens warned against. Consider just two examples. Under normal circumstances, a visual target’s movement always involves a change of that object’s shape and/or position. This confounding of motion and position change has made it hard to connect psychophysical responses to motion alone. To break the confound, Nakayama and Tyler (1981) generated matrices in which black and white cells alternated at random. All the cells in a row were shifted back and forth, left and right; with appropriate rates of shift, observers saw oscillatory motion. The cells were small (<3 minarc), and spatially quite dense. Moreover, because all cells of the same color were indistinguishable from one another, the positions of individual elements could not be tracked. Despite the absence of position information, observers could detect the oscillatory motion generated by shifts of pattern elements. Take a second example of a stimulus designed to test an hypothesis about motion perception. To explore how the visual system combines or segregates spatially-intermingled motions in different directions, Qian, Andersen and Adelson (1994) created displays whose every local region contained balanced, opposite directions of motion. The locally-opposed directions tended to cancel one another, which caused observers to see no overall

motion. This chapter offers numerous other examples of complex stimuli specifically designed to probe particular aspects of motion perception.

Overview of motion processing stages

Where appropriate, this chapter relates psychophysical results on motion perception to underlying neural mechanisms. An interest in establishing such connections drives much contemporary research into visual motion. For this reason, it will be helpful to provide a broad overview of the anatomy and physiology of those portions of the primate visual system explicitly involved in the analysis of motion information (see Figure 2); for a more detailed account, see Croner and Albright (1999).

Among the neurons in the visual systems of primates, cells selectively responsive to direction of motion are first encountered in area V1, the primary visual cortex, which is located in the occipital lobe. Such neurons are often described as “tuned” for direction.³ Beginning with the landmark work of Hubel and Wiesel (1968), it has been known that a significant fraction of V1 neurons respond best when a contour moves through their receptive fields in a particular direction; responses are significantly diminished when movement is in the opposite direction. Different neurons have different preferred directions of motion, with all directions around the clock represented within the ensemble of neurons. This inaugural stage of processing comprises a local analysis of motion energy. In this analysis, direction-selective neurons act as “filters” that register the presence of component features of moving objects within the local regions of their receptive fields (Emerson, Bergen, & Adelson, 1992).

The outputs of these local filters in V1, in turn, activate second-stage analyzers that integrate motion signals over more extended regions of visual space. This second-stage analysis begins with neurons in the middle temporal visual area, or MT as it is typically called. MT receives some of its input directly from V1, and the rest indirectly from V1, via areas V2 and V3. Nearly all neurons in MT are selective for the direction and speed of stimulus motion, again with the range of preferred directions among neurons spanning 360 degrees. MT neurons have larger receptive fields than do V1 neurons, which means they can integrate motion signals over larger regions of visual space. Moreover, a given MT neuron will respond to motion in its preferred direction regardless whether those motion signals are carried by luminance, color or texture. MT neurons, in other words, exhibit form invariance (Croner & Albright, 1999), implying that those neurons register motion information *per se*. MT neurons, in turn, project to higher visual areas that encode more complex forms of motion including expansion and rotation (Tanaka & Saito, 1989) and motion-defined boundaries (Van Oostende, Sunaert, Van Hecke, Marchal, & Orban, 1997). Outputs from MT also make their way to visual areas in the frontal lobe concerned with the control of eye movements (Bichot, Thompson, Chenchal Rao, & Schall, 2001; Schall, 2000).

A great many studies implicate Area MT in the perception of motion. Though neurons in Area MT certainly contribute to the perception of motion, it is clear that this

³ As Parker and Newsome put it (1998, p. 229), “A neuron is considered to be ‘tuned’ if the response is strongest to a particular value (or narrow range of values) of the stimulus and declines monotonically as stimulus values depart from this ‘preferred’ value.”

is not the sole site where neurons extract significant motion information. Actually, various aspects of motion perception depend upon the neural computations carried out in different areas of the cortex. Normally, motion perception depends upon activity distributed over many areas of the brain, each extracting somewhat different information from the retinal image. Complicating matters, in macaque monkeys, creatures with visual system highly similar to *Homo sapiens*, back-projections from MT to V1 have been demonstrated (Beckers & Homberg, 1992). In humans, initial evidence suggests that this back-projection may be important for conscious awareness of visual motion. To explore this idea, Pascual-Leone and Walsh (2001) applied brief pulses of magnetic energy⁴ to spatially-restricted regions of the scalps of human observers. This technique is known as transcranial magnetic stimulation, or TMS. When the localized pulses are adjusted in duration, frequency and amplitude, and are delivered to particular regions of the scalp, TMS creates sensations of flashes of light. Called “phosphenes,” these flashes appear to move when the pulses are delivered to the scalp overlaying visual area MT, but are stationary when TMS is delivered to the scalp that overlays area V1. By applying separate TMS pulses asynchronously to V1 and MT, Pascual-Leone and Walsh obliterated observers’ conscious experience of the moving phosphenes that were ordinarily evoked by MT stimulation. This result required that TMS to V1 be delivered some tens of milliseconds after MT was stimulated. Presumably, the obliteration of motion perception is caused by a disruption of a re-entrant, backprojections from MT to V1. A similar result was reported by Beckers and Homberg (1992).

The preceding description of the motion pathway was based mainly on physiological and anatomical studies of non-human primates. During the past decade, understanding of motion perception’s neuronal substrates in humans has been significantly advanced by the use of brain imaging techniques, primarily functional magnetic resonance imaging (fMRI). This growing literature, has identified at least a dozen distinct regions in which neurons respond to visual motion. These regions in the human brain stretch from the occipital lobe to the frontal lobe (Culham, He, Dukelow, & Verstraten, 2001; Sunaert, Van Hecke, Marchal, & Orban, 1999). Among motion-responsive regions are area V1 (which responds to almost any moving pattern, as well as to stimulus flicker) and the MT/MST complex, located on the brain’s lateral surface near the junction of the occipital, parietal and temporal lobes. This region, which we shall refer to as MT+, responds weakly to flicker but strongly to coherent motion, including optic flow patterns (see below). Other important motion areas include area KO (for kinetic occipital), which responds preferentially to motion-defined boundaries, and area STS (for superior temporal sulcus), which is especially responsive to patterns of motion portraying biological motion. As appropriate, brain imaging results are introduced throughout this chapter to clarify the neural computations that make motion

⁴ TMS offers a powerful tool for investigating cognitive or perceptual neural circuitry (Pascual-Leone, Walsh, & Rothwell, 2000), including circuitry that supports various aspects of motion perception (for example Hotson & Anand, 1999; Walsh, Ellison, Battelli, & Cowey, 1998). When the TMS coil is positioned against an observer’s skull, a powerful, focused magnetic field hits and penetrates the skull. The field penetrates superficial layers of the cerebral cortex, and can temporarily terminate or modify currently ongoing neural activity, or alter neural activity that is about to begin.

perception possible.

With this overview in place, we can now tackle various aspects of motion perception that make it so crucially important for guidance of our everyday activities.

THE LIMITS OF MOTION PERCEPTION

Motion detection

Visual motion can be construed as an event that unfolds over space and time. Distilled to the simplest case, motion involves a continuous change in the spatial position of a single object over time; this can be depicted in the form of a space/time plot in which spatial position along one dimension is plotted as the function of time in Figure 3. Intuitively, one might expect that the ease with which this kind of simple event can be seen would depend on the magnitude of the displacement over time and the rate at which that displacement occurred. And there is truth to this intuition. Consider, for example, the movement of a clock's minute hand. You cannot see the clock hand's gradual progression, but intuitively we know that it has moved because its position has changed over time. Motion perception, however, need not involve any kind of intuitive process; motion is a direct experience, uniquely specified by the visual system (Exner, 1888; Nakayama, 1981; Thorson, Lange, & Biederman-Thorson, 1969). But how sensitive is the system that generates this experience? What is the lower limit for detection of motion? Measurements with a single moving object show that for a sensation of motion to be experienced the object must traverse at least 1 minarc. (This distance approximately the edge-to-edge lateral separation between adjacent alphanumeric characters on this page viewed at arm's length.) This value varies, however, depending on the object's duration, velocity and luminance, as well as with the region of the retina stimulated -- an up-to-date summary of this literature is given by Tayama (2000).

Very seldom, however, are we called upon to detect the motion of a single object appearing in isolation. Instead, most everyday detection of motion involves the detection of one object's (or group of objects') motion relative to another object (or set of objects). And at this challenge --detecting relative motion-- humans excel. We are vastly better at detecting relative motion than we are at detecting absolute motion (for example, Leibowitz, 1955). This is strikingly demonstrated in a study by Lappin et al (2001) in which observers viewed an array of three, horizontally aligned "blobs" (circular, Gaussian luminance patches). In one condition, all three blobs moved laterally back and forth in unison ("in-phase" motion), while in another condition the center blob's direction of motion was opposite to that of the flanking blobs ("anti-phase," or relative motion). Motion acuity -- the smallest detectable motion displacement -- was actually better for the anti-phase condition compared to the in-phase condition. This means, in other words, that tiny motion displacements visible in the anti-phase condition were impossible to see when there was no relative motion. This finding dovetails nicely with earlier results showing that humans are keenly sensitive to shearing motion (Nakayama, 1981; Snowden, 1992), which also entails relative motion, not overall, rigid displacement. This "direction contrast" amplifies differences in motion vectors in neighboring regions of the visual field. Keep in mind that in of the research described below, motion performance is measured for moving targets that appear within some background framework.

Intuition suggests that motion might render an object less detectable. (Think for example, of the difficulty one experiences trying to read a newspaper in a moving

subway.) But under most circumstances exactly the opposite is true. Especially for objects with significant energy at low spatial frequencies motion can render an otherwise invisible object visually conspicuous (Robson, 1966). To experience this for yourself, hold an object between the uniform surface of a wall and a light source in order to create a large, very faint shadow on the wall. When the non-moving dim shadow fades to invisibility, move the occluding object and notice how the shadow abruptly springs into existence. The familiar textbook example of seeing the retinal blood vessels in your own eye by jiggling (i.e., moving) a light source on your sclera is another instance of motion's ability to reveal what otherwise would have been undetectable.

To characterize more precisely the optimal stimulus for the motion system, researchers have exploited the lower envelope principle. Barlow first called attention to the principle forty years ago, and most recently framed it this way: "Sensory thresholds are set by the class of sensory unit that has the lowest threshold for the particular stimulus used and are little influenced by the presence or absence of responses in the enormous number of other neurons that are less sensitive to that stimulus (Barlow, 1995, p. 418)." Discussing various applications of the principle to sensory systems, Parker and Newsome note, "in its pure form, the lower envelope principle means literally that a single neuron governs the behavioral threshold. The development of the lower envelope principle has been very much a reaction by neurophysiologists to the formerly prevalent notion that single neurons are inherently unreliable devices (1998, p. 242)."

In an attempt to apply the lower envelope principle to motion, Watson and Turano (1995) measured the minimum contrast at which observers could discriminate direction of movement. Their test stimuli were drawn from a family of patterns known as Gabor functions. Each Gabor function comprises a sinusoidal grating that has been multiplied, point by point, by the values of a two-dimensional Gaussian function. This multiplication modulates the sinusoid's contrast, producing a pattern whose contrast falls off smoothly in all directions from a maximum at the pattern's center.⁵ In search of the most easily seen moving stimulus, Watson and Turano independently varied the spatiotemporal characteristics of both the sinusoidal grating and its modulating Gaussian function. The stimulus yielding the lowest contrast threshold was a sinusoidal grating with spatial frequency of 3 cycles/degree and drift rate of 5 Hz, with a width and height of 0.44 degrees visual angle and a duration of 0.13 seconds. Applying the lower envelope principle, these parameters describe the most sensitive of the neural filters in the direction-extracting system.

Stimuli like those used by Watson and Turano and others have the advantage of limiting stimulus energy to particular bands of spatial and temporal frequencies. Although such grating patterns are often used to explore motion perception, they bring some disadvantages. For example, they confound motion with orientation, and they can appear to move only in the directions perpendicular to their orientation. To get

⁵ These functions bear the name of Dennis Gabor, a Hungarian engineer and applied physicist. Gabor won the 1971 Nobel Prize in Physics for his work on wavefront reconstruction in optical holography. In vision research, Gabor functions are used as stimuli; also, they are good approximations to many visual neurons' spatiotemporal receptive fields.

around these limitations, researchers devised an entirely different class of motion stimuli, arrays of irregularly spaced, moving elements such as blobs or dots. These computer-generated stimuli are commonly known as random dot cinematograms (RDCs), and there are several different algorithms for generating them. Distilled down to their essence, most RDCs consist of “signal” dots all of which move in a given direction (or within a given range of directions) randomly intermingled with “noise” dots each of which moves in a random direction.⁶ When the proportion of signal dots is high, dots in the RDC appear to move coherently in the general direction of those signal dots; when signal dots comprise only a small fraction of the RDC, the sense of motion coherence is weak or, in the limit, is absent entirely. Motion threshold is defined as the minimum percentage of “signal” dots necessary for detection of coherent motion. It should be stressed that the information supporting detection performance in these stochastic stimuli must be extracted globally: observers cannot perform well simply by attending to a single dot or to a restricted region of the display.

Humans exhibit remarkable sensitivity to coherent motion in RDCs. Under optimal conditions, observers can detect signal percentages as small as 5% (Scase et al, 1996), and this holds for signal dots portraying translational motion, rotational motion and expansion/contraction (but see Ahlström & Börjesson, 1996; Blake & Aiba, 1998). Absolute threshold values vary with display size (Burr, Morrone, & Vaina, 1998), dot density (Watamaniuk, 1993) and exposure duration (Watamaniuk, Sekuler, & Williams, 1989), but not with direction of motion (Gros, Blake, & Hiris, 1998). Visual cues that allow segmentation of signal dots from noise dots -- such as color or contrast polarity -- can substantially enhance detection of motion (Croner & Albright, 1997; but see Heidenreich, 1995). Prior exposure, or adaptation, to strongly coherent motion in a given direction temporarily elevates coherence thresholds for directions of motion within roughly 40 degrees of the adapting direction, with the elevation in threshold being largest at the adapting direction (Heidenreich, 1995). Thresholds are also affected by higher-order variables such as visual attention to a particular direction of motion (Raymond, 2000; Raymond, O'Donnell, & Tipper, 1998), a point discussed in later in this chapter.

In an influential series of experiments, Newsome and colleagues used RDCs to test motion detection in monkeys. They recorded neural responses from single cells in areas MT and MST of the monkey's brain while the monkey tried to detect motion (see review of this work by Parker and Newsome 1998). In general, the monkey's behavioral threshold for detecting motion was very close to the neural threshold derived for some individual directionally-selective neurons. The correspondence between neural thresholds and behavioral thresholds points to a tight linkage between neural activity in MT and MST and the monkey's performance, with “neural” thresholds corresponding closely to behavioral thresholds (Celebrini & Newsome, 1994; Newsome, Britten, & Movshon, 1989).

⁶ The terms “signal” and “noise,” commonplace in psychophysics, derive from engineering and communication sciences. There, the task facing a biological or machine detector is portrayed as the extraction of a message (signal) from a stream of input, some portion of which (noise) is irrelevant or even antithetical to the extraction of the message. These definitions enable investigators to characterize a detector's sensitivity in terms of ratio of signal to noise that just allows signal extraction.

This linkage was further strengthened by the finding that electrical stimulation of neurons in MT (Salzman, Murasugi, Britten, & Newsome, 1992) or in MST (Celebrini & Newsome, 1994) can bias a monkey's perceptual report of motion direction in RDCs. So, direct stimulation of MT neurons tuned to leftward motion increased the probability that the RDC would appear to move in that direction. Direct electrical stimulation of particular clusters of MT neurons, then, was perceptually equivalent to the effect normally produced by an RDC moving in the neurons' preferred direction with a particular coherence level. Lesion studies lend additional support to the idea that MT participates in motion perception. Lesions encompassing areas MT and/or MST produce permanent impairments in the ability to extract coherent motion from RDCs (Newsome & Paré, 1988; Rudolph & Pasternak, 1999). Comparable deficits in motion sensitivity to RDCs are found in human patients with damage to an area of the brain homologous to primate MT (Schenk & Zihl, 1997).

Trajectory detection.

Algorithms that generate RDC stimuli ordinarily prevent any signal dot from moving in a constant direction throughout the entire animation sequence. This is done to prevent observers from basing judgements on any single dot's trajectory, instead forcing judgments to arise from the integration of many motion vectors. In the natural environment, however, there are many instances where it is important to detect one object's movement in the presence of other, distracting moving objects, and, for that matter, despite temporary occlusions in the target object's path. To take an unpleasant example, a lion in hunting mode can visually track the path of one particular zebra in a herd, even when all other members of the herd are moving about in random directions, and even though the target zebra is temporarily obscured by vegetation or other opaque objects. To learn how vision manages such feats, Watamaniuk and colleagues (1995) measured observers' ability to detect the presence of a single dot moving on a fixed, straight path amidst a field of noise dots whose directions changed randomly over time. The signal dot was identical to the noise dots in luminance, size, and speed. For a stimulus duration of 500 msec, motion was detected 90% of the time, even when there were as many as 250 noise dots. Under such conditions, the proportion of signal dots was minute, just 0.4%.

Watamaniuk et al. (1995) also found that a single moving dot's trajectory is easily seen even when that trajectory was interrupted by a series of opaque occluders (Figure 4). Detection of the dot's motion across three path segments, each separated by an occluder one degree wide, was essentially as good as detection of the same motion over a single, uninterrupted path of equivalent length. So the motion signal generated by the moving object was essentially unaffected by temporary interruptions of that signal when the object was occluded. However, when the dot's trajectory was momentarily interrupted by a different kind of occluder, detection of the moving dot fell dramatically. Here, occluding regions were filled with random-direction motions (noise), similar to the random motions in the display's other regions. The dot disappeared when it entered a noisy occluding region and reappeared when it left that region. Within each noisy occluder the virtual path of the signal dot was probably masked or distorted by similarly-directed motion vectors in the noise. These noise directions led the directional signal astray, reducing the precision of matches from one segment of the dot's trajectory to the next. (This is an instance of the so-called motion correspondence problem, which is discussed in a subsequent section.) Because the reduction in detectability persisted even when noise-filled occluders lay in a depth

plane different from the regions containing the trajectory, it seems that trajectory detection operates prior to the assignment of depth to image components.

Grzywacz et al. (1995) proposed a model that can account for many observations on trajectory detection. In their model, local connections among motion mechanisms enhance responses that result when mechanisms are stimulated in sequence, roughly in the direction of their directional tuning. These connections are spatiotemporal analogues to the spatial association fields that have been implicated in contour integration (Field, Hayes, & Hess, 1993; Geisler, Perry, Super, & Gallogly, 2001; Sigman, Cecchi, Gilbert, & Magnasco, 2001). From a perceptual perspective, the connections postulated by Grzywacz promote what Gestalt psychology dubbed “good continuation”: Perception will favor trajectories that are smooth rather than trajectories with large abrupt changes in direction. Such spatiotemporal preferences were demonstrated in Metzger’s (1934) observations with objects moving on independent, but intersecting paths.

Because of the spatiotemporally-tuned local connections in Grzywacz’s model each successively stimulated motion mechanism will produce a larger response than the previously stimulated mechanism. Eventually, the response grows large enough to be accurately detected, even though other motion detectors are responding to the background motion noise. Since local connections involve mechanisms with a range of similar directional tuning (spanning about ± 30 degrees), the model can also account for the detection of curved trajectories as well as straight ones (Grzywacz, Watamaniuk, & McKee, 1995; Vergheze, Watamaniuk, McKee, & Grzywacz, 1999).

Grzywacz et al.’s (1995) trajectory network model postulates facilitation of signals in sequentially-stimulated motion mechanisms, a result observed by Vergheze et al. (1999). They compared the detectability of two sorts of trajectories. In one, the moving elements followed a single, continuous trajectory of length L ; in the other, the same elements traced out n separate trajectories, with a gap between successive segments. Because each segment was L/n long, their summed lengths were the same as the length of the single uninterrupted trajectory. Vergheze and colleagues wanted to know how the motion system would sum the motion signals contained in these segments. For a computational benchmark, they drew upon the notion of probability summation. In its simplest form, probability summation predicts that detection of a signal should be equal to the sum of the square of the number of independent stimulus elements (Graham, 1989; Watson, 1979). In Vergheze et al.’s experiments, the number of independent stimulus elements is n . This computational rule gave a good account of detection when trajectory segments were short enough so that each segment was likely to be detected by just a single mechanism. But with longer trajectories, probability summation failed badly. For example, a single, 200 msec trajectory was about three times more detectable than two spatially isolated 100 msec trajectories presented one after another, which kept L constant. So, the detection of an extended trajectory cannot be explained by activation of a series of independent motion detectors whose outputs are summed linearly. Instead, the result points to significant spatiotemporally-tuned interactions among local motion units, of the kind postulated in Grzywacz et al.’s (1995) trajectory network model.

Motion Discrimination: Direction, Speed and Coherence

Having highlighted key determinants of motion detection, we turn now to motion discrimination. In tasks used to measure discrimination, observers must not only detect the motion (see that motion is present), but must also judge one or more of

that motion's essential characteristics, such as direction, speed or coherence.

In experiments on motion perception sinusoidal grating stimuli were used to investigate discrimination of direction of motion at or near the contrast detection threshold for that moving stimulus. In general, the contrast at which observers could just detect the presence of a moving stimulus was also sufficient to identify its direction of motion (Derrington & Henning, 1993; Levinson & Sekuler, 1975; Watson, Thompson, Murphy, & Nachmias, 1980). Such results suggested that visual mechanisms that extract motion signals are 'labelled' for direction. In this context, a neural mechanism is said to be labelled for some elementary sensation, S , if activity in the mechanism is sufficient to generate the experience of S (Watson & Robson, 1981). Although much has been learned about motion from experiments using grating stimuli, an exclusive reliance on such stimuli encounters substantial limitations. If an observer cannot see the ends of the grating (the usual case in motion experiments), unambiguous motions can only be produced in the two directions orthogonal to the grating's orientation, an ambiguity known as the "aperture problem" (see below). Thus, creating other directions of motion requires a change in the orientation of the grating, confounding direction and orientation. Partly to avoid this potential confound, Westheimer and Wehrhahn (1994) measured direction discrimination for a single moving spot. They found that at high speeds (30 degrees/second), direction discrimination was equal to that for orientation discrimination of a static line of the same length as the distance travelled by the moving spot and presented for the same duration. This suggests that at high speeds, the moving spot created an oriented smear (a virtual line) on which observers could have based their judgements. [Work by Geisler (1999) makes a similar point.] This potential problem is avoided, however, by using RDCs in which the directions of individual dots change frequently and/or if each dot has a limited lifetime.

Williams and colleagues (1984) created RDCs in which individual elements were assigned new directions of motion each frame. When directions were drawn from a distribution of directions spanning 180 degrees or less, observers saw the entire fields of dots move in the direction of the distribution's mean, even though the random movements of individual dots were also visible. This meant that the motion mechanisms responsible for this global motion integrate direction information over much, if not the entire, motion display. Using a RDC in which each dot chose a new direction of motion each frame from a distribution of directions spanning 30 degrees, (Ball, Sekuler, & Machamer, 1983) determined that these motion mechanisms' direction tuning is broad: two directions of motion had to be separated by 120 deg before they no longer stimulated the same mechanism.

With such broadly-tuned motion mechanisms to draw upon, how well can observers discriminate small differences in direction? To answer this question, Watamaniuk and colleagues (1989) measured direction discrimination for fields of random dots. When all dots moved in the same direction, observers reliably discriminated directions that differed by only 1 degree. Moreover, this threshold was relatively resistant to the inclusion of random motions that potentially interfere with signal extraction. When similar measures were made with RDCs whose elements moved in a range of directions, once the range of directions in the RDC exceeded 30 degrees, direction discrimination thresholds increased with direction range. In addition, direction discrimination improved as exposure duration lengthened, up to at least 300-400 msec, and as size of the stimulus increased, up to at least 10 deg diameter (Watamaniuk & Sekuler, 1992; Watamaniuk, Sekuler, & Williams, 1989). So, the motion system is robust in the presence of random motions, and it can also produce precise

discriminations using mechanisms that are broadly tuned.

RDCs can generate motion percepts simultaneously on two different spatial scales. In particular, the perception of global flow can co-exist with the perception of individual dots' small random motions. Watamaniuk et al. (1998) showed that direction information was encoded independently on the two spatial scales, global and local. In their RDCs, a central single dot moved in a constant direction while the remaining 100-150 dots were assigned a new direction of motion each frame from a distribution spanning 160 degrees. The global flow direction and the direction of the constant-direction dot were always similar, but both varied from trial to trial. After a brief presentation of the RDC, a tone told the observer which motion, global or local, was to be judged. Under these conditions, observers judged either direction just as well as if they had been told in advance which direction to judge. This suggests that visual information on different spatial scales is processed simultaneously with little interference.

As indicated earlier, motion detection is isotropic. For direction discrimination, though, performance is anisotropic, varying strongly with the test direction. Discrimination thresholds are lowest (i.e., performance is best) for directions at and near the 'cardinal' directions of up, down, left and right (Gros, Blake, & Hiris, 1998). This "oblique effect" could reflect a disproportion in the number of neurons tuned to cardinal directions; alternatively, it could arise from narrower directional tuning for neurons tuned to cardinal directions. The absence of an oblique effect for detection of motion points to the second interpretation, although the question remains open.

Direction Change

In the natural world, objects often change direction, and responses to such changes can be quite important. For example, direction changes by inanimate objects may result from a collision; direction changes by living creatures may convey biologically-significant information, such as the information produced by a series of hand gestures. Direction-tuned neurons in area MT of primate cortex have been shown to be efficient encoders of such changes (Buracas, Zador, DeWeese, & Albright, 1998). Synchronized changes in direction of motion among an array of small objects promote perceptual grouping of those objects into a coherent shape (Lee & Blake, 1999).

Dzhafarov and colleagues (1993) proposed a formal computational model for responses to changes in speed and direction of motion. The model, which has been extended by Mateeff, Genova and Hohnsbein (1999), incorporated a process that normalized all stimulus velocity signals registered prior to any change.⁷ Velocity normalization is an instance of visual adaptation processes that reduce redundancy in

⁷ Normalization refers to various linear operations that transform some data set, D , into a new set, D' , while preserving particular types of numerical relationships among the set's members. Of two common forms of normalization, subtractive normalization of D can be described as $D' = D - k$; divisive normalization is represented by $D' = D/k$. Subtractive normalization, which is the form used in Dzhafarov et al.'s model, preserves the relative magnitudes of members of D ; divisive normalization, preserves the proportional relationships among members of D . Heeger (1994; Heeger, Simoncelli, & Movshon, 1996) describes another type of normalization, which operates in the visual cortex to partially suppress the responses of individual neurons.

neural responses by minimizing the total neural activity elicited by any stimulus input (Barlow, 1990). For Dzhafarov et al.'s implementation, assume that some stimulus has an initial velocity V_0 , which changes abruptly to a different velocity, V_1 . As a result of normalization, this change from V_0 to V_1 is detected as though the change were the onset of an initial motion with velocity $|V_1 - V_0|$. In other words, the actual value of V_0 is irrelevant; all that matters is the absolute difference between V_1 and V_0 . This basic computation successfully predicted observers' speed of response to change with a considerable variety of values for V_0 and V_1 . Sekuler, Sekuler and Sekuler (1993) examined this normalization process further. In an attempt to disrupt the extraction of information about V_0 , they injected transients such as temporary occlusion or disappearance into the trajectory prior to the change to V_1 . By injecting a transient at various times during and after V_0 , they were able to interrupt or freeze normalization. Even a relatively brief disappearance of the moving target reset normalization entirely, erasing all the velocity information that had been extracted up to that point.

Speed Discrimination

During their daily routine, humans frequently make judgments about the speed of moving objects. Consider, for example, the simple act of judging the speed of nearby cars when changing lanes on the highway. Or running across the lawn to intercept a small child who is crawling toward a street. How do humans judge speed in these kinds of situations? Because objects moving at different speeds cover different distances in any given temporal interval, observers could use distance traveled as a cue to speed. Alternatively, the time needed to travel some criterion distance might also be used as a cue to speed. So how can one measure speed discrimination without confounding influences of time and distance? The usual experimental approach is to randomize presentation time over a range large enough so that duration and, hence, distance cues become unreliable (McKee & Watamaniuk, 1994).

Speed discrimination thresholds are typically presented as Weber fractions ($\Delta V/V$) that specify the proportional difference in speed needed to produce reliable discrimination. The smallest increment in speed that can be reliably detected (ΔV) is divided by the mean or base speed (V). Most studies have reported Weber fractions in the range 0.04-0.08, with various types of stimuli, including moving lines, dot fields, or sinusoidal gratings (Bravo & Watamaniuk, 1995; Brown, 1961; DeBruyn & Orban, 1988; McKee, 1981; Nakayama, 1981; Orban, de Wolf, & Maes, 1984; Pasternak, 1987; Turano & Pantle, 1989; Watamaniuk & Duchon, 1992). The Weber fraction's constancy means that the smallest detectable increment in speed increases with the base or starting speed. Although the Weber fraction for speed discrimination is fairly constant over a variety of test conditions, perceived speed can be altered by any number of stimulus parameters. For example, Katz et al. (1990) reported that briefly-presented drifting stimuli appear to move faster than stimuli presented for longer durations. A grating's drift or movement rate takes deg/second as its units, where deg signifies degrees of visual angle. It's important to distinguish drift rate from a related variable, temporal frequency, which takes units of Hz or, equivalently, cycles/second. The relationship between a grating's drift rate and the temporal frequency produced by that drift takes account of the grating's spatial structure:

$$\text{Temporal frequency (Hz)} = \text{Drift rate (deg/sec)} \times \text{Spatial frequency (cycles/deg)}$$

Perceived speed of movement (perceived drift rate) varies with a grating's spatial frequency, which takes units of cycles/degree. Sinusoidal gratings presented at lower temporal frequencies (and hence lower spatial frequencies) appear slower than gratings moving at the same speed but with higher temporal frequency (Diener, Wist, Dichgans, & Brandt, 1976; McKee, Silverman, & Nakayama, 1986; Smith & Edgar, 1991). Turning to other variables that affect perceived speed, when gratings of different contrast move at the same physical speed, a lower-contrast grating appears to move more slowly (Stone & Thompson, 1992; Thompson, 1982). Furthermore, objects seen in the periphery appear to move more slowly than foveal objects of the same speed (Johnston & Wright, 1986; Tynan & Sekuler, 1982). Finally, the perception of an object's speed is adversely affected at low luminance levels, which correspond to rod-mediated, scotopic vision. Gegenfurtner et al. (2000) showed that a moving object's perceived speed is considerably slowed in rod-mediated vision, compared to its perceived speed in cone-mediated vision. To understand the likely basis for this result, note that differences in rod and cone time constants indicate that rods average information over longer periods of time than do cones. Gegenfurtner and colleagues speculate that the rods' extended temporal averaging attenuates motion signals that would be generated in detectors tuned to high velocities. The reduction in such signals causes the reduction in perceived speed under rod-dominated conditions. Grossman and Blake (1999) found that perception of biological motion and structure from motion, too, were impaired under scotopic viewing conditions. Such findings have clear implications for driving safety at night on poorly illuminated roads.

IDEAL OBSERVERS AND MOTION ANALYSIS.

Measurements long ago established that the visual system was an extraordinarily efficient detector of light (Hecht, Shlaer, & Pirenne, 1942; Rose, 1948). To assess how efficient vision is when it performs other tasks, researchers often turn to ideal-observer models. In its most common form, an ideal-observer comprises a mathematical model of a theoretical observer who has complete and perfect knowledge of all relevant stimulus and task statistics; in addition, this theoretical ideal makes statistically optimal decisions when transforming sensory information into psychophysical responses. Ideal observer models afford interesting benchmarks for the fallibility of human observers, in comparison to the theoretical limit represented by an ideal observer. Ideal observer models have been used to compare human and ideal performance for tasks such as detecting changes in spatial patterns (for example Barlow, 1978; Barlow, 1980; Barlow & Reeves, 1979; Burgess & Barlow, 1983; Burgess, Wagner, Jennings, & Barlow, 1981; van Meeteren & Barlow, 1981).

Random, unpredictable variability in the stimulus limits an ideal observer's performance because it subverts the observer's otherwise perfect knowledge of the stimulus to be detected. Such random variability is known as noise.⁸ Ideal observer

⁸ In virtually any psychophysical experiment, valuable information can be gained from measurements made with stimuli to which various amounts of noise have been added (Pelli & Farell, 1999). Noise can assume different forms, depending upon the observer's task. For example, when the task involves detection of a static form, the noise usually comprises independent, random luminance values added to each element in the

models try to predict how humans and ideal observers might perform when each must extract information from a noisy stimulus. Increasing stimulus noise leads any observer --human as well ideal-- to make more errors, such as failures to detect the stimulus, misclassification of a stimulus, or declaring that a stimulus is present when it actually is not. How closely a human observer approximates the theoretical ideal defines the human observer's efficiency. This statistic is given by the square of the ratio of human performance to ideal performance, where both performance measures are expressed as d' values (Macmillan, in press). Thus, if a human observer's performance were ideal, efficiency would be 1.0. Detailed comparisons of human and ideal observers create valuable, diagnostic opportunities for identifying and quantifying components that limit human performance (Geisler, 1989).

In the first application of ideal observer analysis to visual motion perception, Watamaniuk (1993) devised an ideal observer model that discriminated the direction of global flow generated in random dot cinematograms. In specially-constructed RDCs, the directions in which each dot moved over successive frames were chosen randomly, with replacement, from a Gaussian distribution of directions. The unusual algorithm for generating element movement made the stimulus noisy: the algorithm introduced a random discrepancy between the stimulus' actual directional content and the nominal, or average, directional content represented by the sampled distribution. To vary the magnitude of this discrepancy, in different conditions directions were drawn from one of five Gaussian distributions with different standard deviations. The larger the standard deviation, the greater the mean absolute discrepancy between actual and nominal direction information in the stimulus. This introduced random sampling noise into the stimuli. Because the ideal observer had to rely only on its knowledge of the nominal stimulus, introduction of variability (noise) into the actual stimulus reduced the observer's performance.

Direction discrimination was measured for each direction distribution for a range of stimulus durations, stimulus areas, and spatial densities of dots. As expected, human performance was always poorer than the ideal observer. Efficiency generally decreased with increased stimulus area or density, and remained constant as duration increased from 100-500 msec. However, the data clearly showed that efficiency increased as stimulus noise increased, reaching averaged values of up to 0.35. This suggests that the human visual system was less influenced by the random noise than the ideal observer. Note that high efficiency does not mean high level of performance. Because efficiency is a ratio of human to ideal performance, high efficiency can be obtained at any level of performance. In fact, Watamaniuk found highest efficiency for direction discrimination when average performance was at a d' of about 0.75, which translates into a percent correct discrimination of about 70%. Finally, Watamaniuk identified several factors that undermine human performance, including human vision's limited spatial and temporal summation.

Watamaniuk's ideal observer was designed to discriminate one motion from

stimulus display (Bennett, Sekuler, & Ozin, 1999; Gold, Bennett, & Sekuler, 1999); when the task involves detection of symmetry in a pattern, noise may be introduced by randomly altering the position of each of the pattern's elements (Barlow & Reeves, 1979); or, when the task requires identification of global motion direction in an RDC, noise can be generated by randomizing the directions comprising the RDC (Watamaniuk, 1993).

another, but a comparable ideal observer could be designed for another task, to detect motion. Because the visual system exploits different sources of neural information for direction discrimination and for motion detection (Hol & Treue, 2001), specifying an ideal observer analysis for a new task can be far from a trivial matter. Barlow and Triparthy (1997) applied an ideal observer analysis to the detection of coherent motion embedded in random motion noise. Human as well as ideal observers received 2-alternative forced choice tests. They had to identify which of two intervals contained coherent motion rather than completely random directional noise. Noise was introduced into the stimulus, by making the frame-by-frame positioning of the coherently-moving dots less precise. As precision declined, efficiency increased, reaching values of about 0.3. This result points to human vision's relatively coarse spatial resolution, and shows that adding spatial position noise affects the ideal observer more than the human. This coarse spatial resolution represents a low-pass spatial filtering operation, which, Barlow and Triparthy speculate, enhance sensitivity to naturally-occurring motion.

Extending this analysis of motion detection, Baddeley and Triparthy (1998) examined some limitations that might possibly undermine human observers' performance. Using a novel statistical procedure that examined the frame by frame movements of each dot, Baddeley and Triparthy determined that whereas an ideal observer would base decisions on the motions of all dots in the display, human observers seem to use just a proportion of the dots in the display. The same frame-by-frame analysis allowed them to rule out other possible limiting factors, including the idea that human observers differentially weighted directional information generated at various locations in the visual field.

OPTIC FLOW AND STRUCTURE FROM MOTION

Motion affords a potentially powerful information about the 3-dimensional shapes of moving objects as well as about an observer's own movements within the environment populated by those objects. In particular, the movements of objects within an environment create spatiotemporal changes in the light distribution on the retina of a stationary observer. Likewise, an observer's own movements through a stationary environment change his or her retinal image. Such spatiotemporal changes, whatever their origin, are termed optic flow, and they constitute significant sources of visual information. For example, optic flow provides information about the speed, direction, and path of an observer's movements; it can also provide information about the three-dimensional structure of the environment (Koenderink, 1986). In the natural environment, an otherwise camouflaged object -- such as a edible insect -- stands out conspicuously when it moves relative to its background. Any creature blessed with the neural machinery to extract and register the presence and shape of this "target" can secure a meal, courtesy of visual motion.

Historically, research on optic flow has tended to bifurcate into distinct branches. One branch has focused on optic flow's use in steering an observer's movement and heading; the other branch has focused on optic flow's use in revealing the shape and structure of moving objects. In discussing the uses of optic flow information we will respect this bifurcation, although it is worth remembering that both these uses of optic flow arise ultimately from the same source.

Optic flow supports perception of heading.

Conventionally, the patterns of retinal image flow produced by self-motion are represented by an instantaneous velocity field, as illustrated for simple translatory movement in the top panel of Figure 5. Each vector signifies the velocity (direction and speed) in the retinal image of an environmental element. For the case illustrated, the observer's gaze is assumed to coincide with the direction in which the observer is moving. This sets up a radial pattern of optic flow in which the focus of the flow corresponds to the observer's heading. Note that the representation contains no information about the elements' physical attributes, such as their color, shape or size. Instead, they are treated as uniform entities, known as "tokens." Figure 5's lower panel represents the velocity field resulting from an observer's circular movement parallel to the groundplane. This velocity field would be generated, for example, on the retina of an driver whose automobile made a smooth turn.

Although snapshot representations like those in Figure 5 omit information such as acceleration-produced changes or temporary occlusions (Warren, Blackwell, Kurtz, Hatsopoulos, & Kalish, 1991), they are still useful in understanding optic flow's possible role in steering and wayfinding. Koenderink (1986) provides a thorough, accessible mathematical account of the optic flow dynamics that result from various basic types of observer movement. Transformations generated by such movement, no matter how complex, can be decomposed into four basic components: translation, isotropic expansion, rigid rotation, and shear. Summing these basic components in varying amounts can reproduce the original complex optic flow field. As a result, visual mechanisms specialized for extracting these basic components could generate signals that, in aggregate, would represent the complex flow field and, therefore, an observer's movements. The receptive field properties of neurons in MST make those neurons well-suited for extracting optic flow information related to an observer's heading. Such neurons' participation in navigation has been confirmed empirically. Direct electrical stimulation of local clusters of neurons in MST of monkeys altered the monkeys' judgments of heading direction (Britten & van Wezel, 1998). With human observers and human brains, functional neuroimaging reveals large areas of MT+ that respond to particular components of optic flow --such as either circular or radial motion-- but not to simple translation motions in a fronto-parallel plane (Morrone, Tosetti, Montanaro, Fiorentini, Cioni, & Burr, 2000).

Research into optic flow's role in wayfinding has addressed two key issues. First, psychophysical experiments had to determine whether observers can extract heading information from displays containing only optic flow information. Second, knowing that such extraction was indeed feasible, researchers have tried to identify conditions, if any, under which such optic flow information is actually used. We'll consider these two lines of investigation in turn.

Accuracy of heading judgments based on optic flow. For modes of human locomotion such as walking, running, skiing and driving, Cutting (1986, p. 277-279) estimated the accuracy with which people needed to assess their own heading in order to avoid obstacles. For example, suppose a person 1.75 m tall and weighing 65 kg, runs steadily at 5 m/sec (about 11 mph). Assuming that the runner has normal reaction time, in order to swerve to avoid an obstacle such as a tree, the runner would have to judge his or her own heading to an accuracy of at least 1.6 degrees visual angle. If the same person walked rather than ran, far less precise heading judgments would still afford a margin of safety. For downhill skiing, where velocities may be nearly 3 times that of running, the margin of safety narrows to 0.78 degrees visual angle, which heightens

skiing's potential risk.

How do such theoretical estimates square with what human observers can actually do? Empirical measurements of such judgments showed that with paths generated by pure translation across a ground plane, human observers achieve the performance level Cutting stipulated (Warren, Blackwell, & Morris, 1988). Warren and colleagues created 2-dimensional displays in which movements of display elements simulated an optic flow field produced by an observer's movement. Viewing random-dot displays that represent various directions of self-movement through the environment observers can judge that direction to an accuracy of about 1 degree of visual angle (Warren, Blackwell, & Morris, 1988). This level of visual precision is maintained whether the observer's gaze is fixed or shifts with eye movements of the sort that observers might make while walking through a real environment. Furthermore, direction judgements are relatively robust in the face of visual noise added to the otherwise consistent motions of simulated flow fields (van den Berg, 1992).

Most experiments on the heading judgment accuracy have used displays that simulate retinal consequences of simple, straight paths through the environment. In an important departure from this approach, Warren and colleagues (1991) simulated self-motion on a curved path. After viewing random dot fields that simulated some curved path through the environment, observers were shown a distant target and had to judge whether the path they had seen would have taken them to the left or to the right of the target. Warren et al. measured heading thresholds for paths with varying radii of curvature and densities of dots in the flow field. Paths with typical curvature supported thresholds of 1.5 degrees; performance was unchanged by even dramatic reductions in the number of random elements in the display. This finding carries theoretical weight. Neural network-based explorations predicted that a pair of elements seen in two successive views would provide visual information sufficient to support perception of an observer's curved path (Warren, Blackwell, Kurtz, Hatsopoulos, & Kalish, 1991). As a public service for drivers and downhill skiers, we note that accuracy of path judgments is severely degraded at small radii of curvatures (sharper turns).

Under many conditions, then, observers can judge their own paths quite accurately based on optic flow. But examination of this ability has been limited to movement over single, simple paths. Does optic flow play a role when people attempt to navigate complex, multi-legged paths? To answer this question, Kirschen and colleagues (2000) asked people to navigate computer-generated synthetic environments with and without salient optic flow. Using the arrow keys on a computer keyboard to control their own simulated self-movement, Kirschen's participants made repeated trips over what had originally been a novel path. Trips grew faster as participants learned the environment's layout. Because these test environments comprised series of identically textured virtual corridors and intersections, in order to perform the task, participants needed to construct some mental representation of the environment. By varying the rate at which the display was updated, the researchers created optic flow that was either smooth or very choppy. The choppy condition created the impression of a series of separate still views of the environment. The availability of smooth optic flow promoted faster learning of complex paths, mainly by preventing disorientation and backtracking. In a second experiment, participants navigated within a virtual city-block environment, experiencing two different kinds of optic flow as they went. Smooth optic flow enhanced observers' ability to navigate accurately to the remembered position of target objects. Kirschen et al. concluded that when other cues, such as distinctive landmarks, are not available, optic flow can be a significant aid to

navigation.

Mere availability of reliable optic flow information does not guarantee that all observers will be equally able to exploit such information. Here we consider two classes of observers for whom the quality of heading judgments is diminished.

Ball and Sekuler (1986) showed that healthy older people in their 70's and 80's had elevated direction discrimination thresholds. Although these measurements involved random dot motion in a frontoparallel plane rather than heading-related judgments, the older observers' elevated thresholds might be a sign of some general age-related decline in motion perception, particularly perception of multi-element displays and tasks requiring heading judgments. Warren and colleagues (1988) measured heading thresholds for two groups of people, young (mean age about 20 years) and old (mean age of late 60's). Observers saw optical velocity fields that would be produced by observer translation or by observer movement along a curved path. After each display, which lasted about 4 seconds, a vertical line was presented, and observers had to judge whether the heading direction that they had seen would have taken them to the left or right of the line. With straight paths, young observers' heading thresholds averaged 1.1 degrees visual angle, while old observers' heading thresholds were significantly higher, 1.9 degrees of visual angle. With curved paths, the corresponding threshold values averaged 1.4 and 2.9 degrees visual angle. After ruling out ocular and other peripheral causes, Warren et al. suggested that these substantial age-related declines in heading acuity resulted from changes in high level visual processing.

Patients with Alzheimer's disease, a progressive, degenerative disease of the brain, often have difficulty finding their way around their surroundings, even when those surroundings are familiar. Two groups of researchers have connected this difficulty to subnormal processing of information contained in optic flow. If optic flow information actually guides navigation, at least under some circumstances, failure to process that information fully could produce spatial confusion and loss of one's bearings. Rizzo and Nawrot (1998) showed that patients with mild to moderate Alzheimer's disease have particular trouble extracting form or shape information from stimuli in which the form is defined by movement (shape from motion). Moreover, Tetewsky and Duffy showed that some patients with Alzheimer's disease are impaired in extracting directional information from optic flow stimuli (1999). Many of these same patients showed correspondingly poor performance on a test of spatial navigation (wayfinding) ability.

Is optic flow normally used to guide locomotion? Observers' ability to exploit the information in optic flow in order to judge heading does not guarantee that such ability is actually used to control locomotion. To see whether the information was used, Warren et al. (2001) allowed individuals to walk freely in a very large room while wearing a head-mounted display that afforded wide-field stereoscopic vision of computer-generated imagery. At the same time, head position was tracked, and head-position signals were used to update the computer imagery at 60 Hz. While viewing several different sorts of display images, observers attempted to walk as quickly as possible to a visible target such as vertical line or a doorway. Under normal viewing conditions, outside the laboratory, walkers could base their locomotion on information other than the optic flow generated by their own movements. So, a walker could register the position of a seen target in egocentric coordinates and then walk toward that target's position, trying to remain centered on the target. Ordinarily, the optic flow

generated by this egocentering strategy would be identical to the flow generated if an observer used the flow itself to control locomotion. Warren and colleagues broke this normal correlation by feeding participants optic flow information that deviated systematically from what their movements alone would have generated. In a control condition, when no optic flow information was provided, observers' paths indicated they had walked in the egocentric direction of the target. Thus, they homed in on the target by keeping it centered with respect to their bodies. In other conditions, optic flow was introduced, with the unusual arrangement that the focus of expansion was offset from the walking direction (by 10 degrees visual angle). As additional optic flow was introduced into the displays --by adding texture to floors and ceilings -- observers' navigation behavior changed dramatically. Now instead of walking toward the target, as they had been instructed to, they tended to follow the optic flow, which was purposely misleading. It appears then that under normal conditions the visual system depends upon both optic flow and egocentric localization, using the two sources of information in a complementary fashion. Warren and colleagues note that when flow is reduced, say on a grass lawn or at night, reliance on flow is reduced, but in environments that afford considerable optic flow, forested areas for example, locomotor behavior is likely to be dominated by flow.

Optic flow supports collision avoidance.

When an observer is on a collision course with an object, the object generates a characteristic spatiotemporal expansion on the observer's retina. This fact holds equally well, of course, for a moving observer and stationary object, or vice versa. If the rate of movement is constant, the retinal angle subtended by the object grows non-linearly, in a nearly exponential fashion.

Animals as dissimilar as fiddler crabs, chicks, monkeys and human infants, all try to avoid looming patterns created artificially, on television screens (Schiff, 1965). This is true even for newborn infants who have never before encountered a looming stimulus. Apparently, learning plays little role in this behavior.

The rate of retinal image expansion specifies the time to collision --the moment at which a moving observer would reach the object. If at time t an observer D meters away from the object starts moving steadily toward the object, time to collision is given by travel distance divided by travel rate, D/R . Recognizing that solving for time to collision would require information about both D and R , Lee (1980) suggested that vision exploited another, dynamic source of information about time to arrival -- which would make it unnecessary to know either D or R .

Tau, the variable Lee introduced, is the ratio between the current retinal image size and the rate of change in that image size. If the visual system computed a value approximating tau, time to collision would be given by tau's reciprocal. Note this calculation's result is independent of object size.

The connection between imminent collision and rate of expansion of retinal image size holds for any moving creature that has its eyes open -- including birds that fly along and then dive at great speed in order to catch a fish in the water below. Among the best studied of these diving birds is the gannet, a large, web-footed seabird with a sharply-pointed beak. Gannets fly along, wings spread wide until, just before their diving bodies would hit the water's surface, they tuck their wings tight to their sides. Timing is everything. If a gannet performs this maneuver too late, the impact with the water could be quite damaging; if a gannet folds its wings prematurely, its body will be buffeted by cross-winds that will alter the point of entry into the water.

Although we have no wings that need folding, the human visual system carries out similar complex computations. For example, information about rate of expansion can be used to control the braking of an automobile (Lee, 1976), the split-second changes in gait needed when running across rough terrain (Warren, Young, & Lee, 1986), or the various movements and adjustments of the hand required to catch a ball (Savelsbergh, Whiting, & Bootsma, 1991). This coupling between optical expansion and action is not performed consciously. People succeed in these tasks despite being unaware --or being unable to articulate-- what they are doing (Savelsbergh, Whiting, & Bootsma, 1991).

Although tau may be helpful in many circumstances in which people must recognize collision time, it cannot be the only effective cue to collision. In fact, tau would fail under a number of conditions (Tresilian, 1999). For example, if the approach rate is not constant, tau evaluated at any single moment fails to give the correct collision time. Gravity-induced acceleration also undermines tau's usefulness as a predictor of when a falling body will strike some surface. Additionally, with a very slow approach to an object, the rate of image expansion could become so small as to drop below threshold.

Tresilian (1999) discusses tau's limitations, and catalogues other cues that observers could and, in fact, do use. Human observers seem to be quite flexible in making use of available information to solve the collision problem. Thus, Schrater, Knill and Simoncelli (2001) showed that observers can estimate expansion rate in the absence of optic-flow information, using just gradual changes in the scale of random texture elements. In a simulated ball-hitting task, Smith and colleagues (2001) demonstrated that observers optimize performance by adjusting the relative weights given to cues such as the approaching object's angular subtense and its rate of expansion.

It is both plausible and theoretically attractive to postulate that observers adjust cue-weights to match task demands, and to reflect the reliability and availability of various cues. But this theoretical proposition raises a series of theoretical questions about the control architecture that might be used to integrate task-specific processes (Hildreth, 2001). For example, do observers always generate the values of all potential information sources, and then use an optimization strategy to select the most reliable? Or, do observers actually generate only one or two pre-selected weights, based on the characteristics of the situation and task? Such questions must be asked and answered in order to clarify the real value of postulating observers's flexibility in choice of strategy.

Optic flow supports perception of object structure.

Kinetic shape In the laboratory, motion is a potent specifier of shape. This potency was demonstrated in studies by Regan (1989), who created displays in which alphanumeric characters were defined by clusters of dots moving in different directions. To illustrate, imagine a dense array of tiny dots each of which moves. Now suppose a subset of those dots, which fall within the boundaries of a virtual shape, all move in the same direction, while dots outside of the shape's boundaries move in a different direction. (It's important to realize that the region of the virtual shape itself is not necessarily moving; only the dots defining that area move.) People readily see a figure defined by those common motion vectors, and they can judge with excellent accuracy the shape of that figure. Called "kinetic form perception," this ability is conserved in the face of substantial amounts of random motion noise.

The perception of biological motion provides another compelling example of vision's ability to recover object information from motion. When an animal's body moves, the body deforms, that is various parts of the body move relative to one another.

These characteristic relative movements are signatures of normal, biological motion. In fact, when a human moves with stiff or locked body joints, reducing the normal movement of body parts relative to one another, the body's movement looks unnatural, artificial. Because normal biological motion involves deformation of the body, biological motion is classified as nonrigid motion. Although there are many non-biological sources of non-rigid motion --such as the movement of a flag waving in the breeze-- but perception of human body movement has drawn the most interest and research.

In studies of biological motion, the activity and identity of an animate creature are quickly and compellingly extracted from just a dozen or so "light points" placed strategically on the creature's body (Johansson, 1973). In animation sequences that represent the points' movements over time, no single point conveys information about the object or event being depicted. Individual points merely undergo translational and/or elliptical motions. The lights promote perception best when they are placed on joints, the parts of the body whose movements are most diagnostic. Perception of a biological organism that is engaged in an activity requires global integration of motion signals over space and time. As a result, the perception of such animation sequences is literally the creation of motion information (in the same sense that perception of an object in a random-point stereogram is the creation of binocular disparity information). Even a brief view of a point-light display allows an observer to identify the gender of the person in the display (e.g. Kozlowski & Cutting, 1977; Mather & Murdoch, 1994), the activity in which he/she is engaged (Johansson, 1973), or the person's emotional state (e.g. Brownlow, Dixon, Egbert, & Radcliffe, 1997). Human infants as young as 3 months of age can perceive biological motion (e.g. Fox & McDaniel, 1982), as can cats (Blake, 1993). Perception of biological motion is remarkably robust, in that observers can readily discriminate biological from non-biological motion even when the points' contrast changes randomly over time, or when the points are defined entirely by texture and not luminance (Ahlström, Blake, & Ahlström, 1997).

The perception of biological motion may be mediated, at least in part, by unique motion mechanisms. For one thing, information specifying biological motion can be summed over much longer temporal intervals than can that for simple translational motion (Neri, Morrone, & Burr, 1998). For another, damage to specific regions of the brain can impair perception of biological motion while leaving intact other forms of motion perception (Cowey & Vaina, 2000). Conversely, damage to other regions of the brain impair perception of translational motion but have no influence on perception of biological motion (Vaina, Lemay, Bienfang, Choi, & Nakayama, 1990). Finally, functional neuroimaging (Grossman, Donnelly, Price, Morgan, Pickens, Neighbor, & Blake, 2000) has identified regions of the brain, located on the posterior superior temporal sulcus (STS), that are active during viewing biological motion, but not during viewing of the same local motion vectors scrambled in space. These specialized regions are located anterior and superior to MT+.

Kinetic depth In addition to object shape, motion information can also convey the 3D spatial structure of surfaces. A single view of a stationary two-dimensional projection of a three-dimensional object affords little unambiguous depth information. But when the object is set into motion, its projection can produce a clear impression of the object's depth and spatial structure. Wallach and O'Connell (1953) provided an early report of this phenomenon, which they dubbed "the kinetic depth effect." Today, the preferred term for the general class of phenomena is "structure-from-motion," which

encompasses not only the emergence of depth from motion, but also the generation of surfaces and other object related properties. Specially-constructed motion displays have been crucial to understanding the computational and neural basis of structure from motion. Figure 6 illustrates the construction of one such stimulus display. A computer is programmed to create a flat, 2-D projection of a bespeckled, revolving transparent, vertical cylinder. Viewing this 2-D projection, an observer sees intermingled rightward-moving and leftward-moving dots (arising from front and rear surfaces of the cylinder). Additionally, the speed of individual dots will vary systematically across the display, with the fastest moving elements being projections of dots lying closest to the left and right edges of the cylinder's projection. This 2-D projection typically produces a strong percept of structure from motion, giving rise to a bistable percept: a cylinder whose direction of rotates alternates between front leftward and front rightward (Andersen & Bradley, 1998).

In humans, the ability to exploit motion as source of 3-D shape information seems to develop very early. Using an habituation technique, Arterberry and Yonas (2000) showed that infants as young as 8 weeks can perceive 3-dimensional object shape defined just by optic flow. This early access to structure from motion is fortunate because mechanisms that extract static form information are relatively immature in early infancy.

Having seen the potential importance of shape from motion, we can ask how that information is actually extracted from the optic array. Originally, it was thought that 3-D structure in displays like the one in Figure 6 resulted from a series of matches, over successive samples, of the spatial locations at which elements are located (Ullman, 1984). Such an algorithm, in which spatial positions were matched, ignored contributions from local motions, whose central role in structure from motion has been proven conclusively (Andersen & Bradley, 1998).

To explain the neural basis for structure from motion, Nawrot and Blake (1991b) devised a neural network incorporating interactions between stereopsis and direction selective neurons (1991a). Because some neurons in MT are tuned not only for motion, but also for stereo depth, their model incorporated units tuned to combinations of disparity (e.g., near vs. far) and direction of motion (e. g., left vs right). The display shown in Figure 6 would activate two different types of units in a network of neurons: units sensitive to the conjunction of near disparity and rightward motion, and units sensitive to far disparity and leftward motion. Mutual inhibition among units tuned to a given depth plane put them into an opponent relationship, with activation of one type of unit, say "near-depth leftward", tending to reduce responses in units signalling the opposite direction of motion, in this case "near-depth rightward." These within-plane interactions insured that activity at a given depth plane was associated with one direction of motion only. In addition, units tuned to the same direction of motion in near and in far depth planes also exerted mutual inhibition on one another. These between-plane interactions promoted segregation of different directions of motion into different depth planes. Together, these opponent arrangements keep the network from generating contradictory "percepts" within any local area, just as in the physical world, any single, local region on the cylinder's surface cannot simultaneously move leftward and rightward.

Nawrot and Blake assigned their units non-linear stimulus-response activation functions. This non-linearity, together with excitatory connections among similarly-tuned units, allows structure from motion to build up over time, which is consistent with human observers' experiences with many such displays. For example, under

many conditions, the full structure of a display like the cylinder in Figure 6 can take as long as a second or more to emerge fully. Similar, cooperative interactions among like-tuned neural units have been implicated in other aspects of motion perception (Nawrot & Sekuler, 1990; Williams, Phillips, & Sekuler, 1986). Finally, the model's combination of noise (moment to moment variability in response) and opponent organization caused it to reproduce one other notable feature of human perception of structure from motion: a perceptual bistability produced by ambiguous displays like the cylinder. Confirming this general idea, analogous bistability can be seen in the behavior of MT neurons in the brains of awake monkeys who view ambiguous 2-D displays like the projection of the cylinder (Andersen & Bradley, 1998).

Nawrot and Blake's model makes some interesting predictions. First, because kinetic depth and stereo-depth are computed within the same neural network, the model predicts that it should be possible to create stimulus conditions where the two forms of depth -- kinetic and stereoscopic -- will produce equivalent perceptual experiences. That prediction was confirmed in a series of experiments (Nawrot & Blake, 1993). In one experiment, observers viewed two successive animations depicting a rotating sphere. In one display, the impression of a 3D sphere was created solely by kinetic depth; in the other display, retinal disparity was used to portray the 3D sphere. For some disparities, these two displays were indistinguishable under forced-choice testing. A second, related experiment revealed that a weak sense of depth from retinal disparity could be reinforced or, alternatively, could be canceled by a strong kinetic depth stimulus which itself contained no explicit disparity information. This finding, too, points to a common neural substrate for kinetic depth and dynamic stereopsis. A kinetic depth stimulus' capacity to bias stereoscopic depth is comparable to the effect of direct, localized electrical stimulation of direction-selective MT neurons. In awake, behaving monkeys, stimulation of MT has been shown to bias perceptual judgments of depth (DeAngelis, Cumming, & Newsome, 1998).

MOTION TRANSPARENCY

When an observer views a display in which two or more velocities are spatially intermingled, the percept can be one of transparency (two or more separate patterns moving through one another) or coherence (the patterns cohere and move in a single, common direction. Gibson (1959) reported that observers could detect two overlapping planes in a display in which the planes moved at different speeds. Subsequently, Andersen (1989) demonstrated that observers could accurately identify the presence of up to three planes with a similar display with a duration of 2 seconds. When duration is reduced to only 250 msec, observers can distinguish between a display comprising as many as five transparent sheets of dots moving in different directions from a display of dynamic random noise (Mulligan, 1992). And with translatory movements, the process of segregating a display into different planes is fast: With just a 60-msec exposure, observers can correctly identify a 2-plane stimulus. However, if more complex motion patterns such as expansion/contraction and rotations are superimposed and presented for brief durations (85 msec followed by a random noise mask), observers can not accurately identify the component motions (De Bruyn & Orban, 1993).

In transparent displays, the dots defining each separate plane usually move in phase, as an intact pattern in a given direction and speed. But transparency can also be observed if individual dots in the display alternate between two different speeds (Bravo & Watamaniuk, 1995). In these cases, transparency is perceived, as evidenced by

observers' abilities to judge accurately the speeds of individual component motions.

Although two spatially-superimposed sets of dots moving in different directions can generate perceptual transparency, the two sets of motion signals do not operate without mutual interactions. Specifically, the perceived directions of the two sets of dots appear to be pushed apart further than they actually are, a phenomenon referred to as motion repulsion (e.g., Hiris & Blake, 1996; Marshak & Sekuler, 1979). If the superimposed sets of dots' directions differ by less than about 90 deg, the perceived direction of each motion will be pushed away from the direction of the other motion. The magnitude of this repulsion effect depends upon the component motions' speed and upon the density of moving elements (Dakin & Mareschal, 2000; see also Lindsey, 2001). In addition, Snowden (1990) and Verstraten and colleagues (1996) showed that sensitivity to the direction of each motion plane decreases as the speed of the two motion planes become more similar. Thus, although the percept of transparency shows that vision segregates component motions from one another, when the two motion signals are sufficiently similar, each can influence the perception of the other.

The percept of transparency requires that a single local region contain more than a single direction of motion. But what is the size of such transparency-producing regions? Qian, Andersen and Adelson (1994) addressed this issue by generating direction-balanced displays in which dots moving in opposite directions (left and right) were either 'paired,' so that a dot from each set was in close spatial proximity to the other, or 'unpaired'. The observer's task was to judge the degree of transparency in the display. When the paths of the briefly presented dot pairs crossed or when they were separated vertically by 0.2 deg or less, the percept of transparency was abolished. In both cases in which transparency is abolished, the oppositely-moving dots were close to one another. A similar lack of transparency is perceived if two oppositely-moving sine wave gratings are superimposed.

Transparency could be restored to dot and grating displays if the components are sufficiently separated in depth, or if component gratings differ in spatial frequency by about two octaves.⁹ Curran and Braddick (2000) refined this work, showing that if paired dots moved in directions separated by just 60-120 degrees, rather than in opposite directions, the percept was that of coherent global flow in the direction of the direction of the dots' vector average. Presumably, paired directions differing by only 60-120 degrees would not trigger the inhibition that would be generated had the directions been opposite one another, that is if they differed by 180 degrees. Similarly, Lindsey and Todd (1998) found that motion signals embedded in random noise were more easily detected when the component motions moved at right angles to one another, rather than in opposite directions. These data are consistent with the existence of a suppressive stage of motion processing in which oppositely-tuned motion detectors inhibit each other locally. When the directions of spatially-proximate motions are opposite one another, directional signals are perfectly balanced, and mutual, direction-selective inhibition results in no net perceived motion (Qian, Andersen, & Adelson, 1994). When spatially-proximate signals differ by just 90 degrees, the absence of directional inhibition allows the motion to be easily seen, as Lindsey and Todd found.

⁹ A difference of 2 octaves means that the spatial frequencies differ by a factor of four. With a difference this large, the two gratings will be initially registered by different spatial-frequency tuned visual mechanisms (Graham, 1989).

Neurophysiology provides further support for this hypothesis. Qian & Andersen (1994) recorded from V1 and MT neurons during presentation of 'paired' or 'unpaired' motion stimuli and found that V1 cells responded equally well to both types of stimuli, but that MT cells responded better to stimuli in which local motions were 'unpaired.' Snowden et al. (1991) reported data consistent with these results. But what about the brain of humans? Does the human motion complex (MT+) behave as its non-human primate homologue do? Does MT+ exhibit motion opponency? To answer this, Heeger and colleagues (1999) compared multiunit recordings from neurons in monkey cortex to the fMRI activation patterns generated in the human brain. A multiunit recording collects responses from local aggregates of neurons, rather than from individual neurons in isolation. This approach was chosen in order to enhance comparability with fMRI measurements, which represent aggregated neural activity. Heeger and colleagues used equivalent stimuli for both measures of neural activity, and with both species. In humans as well as monkeys, primary visual cortex showed no evidence of motion opponency -- responses to paired and unpaired stimuli (dot patterns or gratings) were essentially the same; however, Area MT in the monkey and its homologue in humans showed considerable motion opponency. In particular, the fMRI activation in MT+ was far stronger with non-paired dot stimuli than with dot patterns in which local, opposite motions were pitted against one another. Thus area MT+ is a site at which direction opponency could initiate the assessment of transparency.

HOW THE VISUAL SYSTEM MEASURES MOTION: THREE PROBLEMS TO SOLVE

This section provides an overview of three major problems that vision must solve in order to provide observers with useful accounts of the motion in the visual field.

Direction selectivity/Reichardt detectors

The first computational account of motion perception arose five decades ago, from the collaboration of Bernhard Hassenstein, a physicist, and Werner Reichardt, a biologist (Borst, 2000). Their product was a simple multiplicative correlation detector made up of two, oppositely-tuned subunits. To understand the detector's operation, imagine that a spot light moves across the retina, successively stimulating different groups of adjacent photoreceptors one after another. To simplify, assume that the moving spot's direction caused the spot to fall first on photoreceptor **A**, and then, after some delay, t , on photoreceptor **B**. As a result, the luminance signal elicited from **A** precedes the signal generated from **B** by t . This delay depends upon two variables, the spatial separation between **A** and **B**, and the speed with which the spot moves. Now to the detector's circuitry. For one of the detector's subunits the luminance signal generated in photoreceptor **A** is multiplied by a delayed luminance signal from a second, neighboring photoreceptor set, **B**. This basic operation is replicated in the detector's other subunit, but in mirror-symmetrical fashion: the two photoreceptors are flip-flopped, and the delay is now applied to the signal from the previously non-delayed photoreceptor. Because of the delays, a spot reaching first **A** and then **B**, generates a larger response in the second subunit than in the first; the same spot travelling at the same speed, but in the opposite direction generates a larger response in the first subunit than in the second. In other words, the numerical difference between the two subunits' responses is directionally selective: motion in one direction generates a positive difference, motion in the opposite direction a negative difference.

The model's simple circuit guarantees that motion sensitivity will reflect a

stimulus' temporal and spatial parameters, which is certainly true of vision. In its first tests, the model was applied to insect vision, exploiting as a behavioral index of motion perception the optomotor reflex of the beetle *Chlorophanus*. The model's success inspired a good deal of research, including work on higher animals. It also promoted the creation of other models that performed a similar computation using different circuitry (Sperling & Lu, 1998).

Although the Hassenstein-Reichardt motion circuit has many virtues, it also has one property that could be considered a flaw: it fails to distinguish between two classes of stimuli that physically are quite different from one another. In particular, the circuit would give equivalent responses to i) a spot that moved smoothly with the proper velocity from the receptive fields of one subunit's receptors to the receptive fields of the other subunit's receptors, and ii) a spot that was presented to one set of receptive fields, then extinguished, and after a delay, presented to the other receptive fields. With proper delays between presentations of the spot, this latter, sampled or stroboscopic motion stimulus, would be indistinguishable from its smooth counterpart. Scaled up to an entire human visual system, this perceptual "error" becomes quite important: it allows the sampled images that comprise film and video sequences to mimic smooth motion. The result of such a sampling process is known as "apparent motion," a designation meant to contrast with smooth or "real" motion. The quality of apparent motion --how smooth the motion appears-- varies with a number of parameters, particularly the rate at which the stimulus is sampled in both space and time domains. As the interval lengthens between successive frames of display, the sampling rate is said to decrease. Intuitively, as sampling rate increases --and successive frames come closer together in time-- the appearance of the sampled stimulus approaches that of a smoothly-moving stimulus.

Watson, Ahumada and Farrell (1986) developed a simple model that predicts whether any spatial and temporal sampling rate would or would not produce the appearance of smooth motion. Their model defines a spatiotemporal range of each observer's "window of visibility." The boundaries of this window, a region in joint spatial and temporal frequency space, define the spatial and temporal frequency limits of the observer's sensitivity to energy in the stimulus. When the stimulus is sampled in time, as for video or film or computer displays, the sampling process generates energy at temporal frequencies in addition to the fundamental frequency. A low sampling rate produces energy over a range of low temporal frequencies; a high sampling rate produces energy over a range of high temporal frequencies. As a result, the higher the sampling rate, the more likely it is that the resulting energy will fall outside the window of visibility, which renders them invisible and perceptually inconsequential. So two stimuli, one smoothly moving and the other representing sampled motion, will appear identical if their spectra within the window of visibility are identical; portions of their spectra that lie outside the window are irrelevant. Using two different strategies for sampling stimuli, Watson and colleagues confirmed the essential validity of their elegantly simple model.

Following Hassenstein and Reichardt, most studies of motion perception have examined responses to drifting modulations of luminance (or chromatic contrast). These stimuli, termed first-order stimuli or Fourier stimuli, would evoke responses in visual mechanisms responsive to spatiotemporal variation in luminance or chromatic contrast. Such stimuli correspond to a dominant species of spatiotemporal modulation encountered everyday, but such stimuli do not exhaust the possibilities. Some stimuli, termed second-order or non-Fourier stimuli, would elude detection by such

mechanisms (Chubb & Sperling, 1988; Pantle, 1973). Nevertheless, perceptually, such stimuli elicit strong motion responses. It is worth emphasizing that second order motion is not merely a creation of the laboratory. A well-camouflaged creature moving against a background with the same texture markings as the creature's own, will generate second order motion only. The same is true for the waving of tree branches when wind blows through a forest, or wheat stalks waving in a wheat field.

Many psychophysical studies comparing first and second order motion have demonstrated clear distinctions between the two. The distinction between the two classes of motion stimuli gains increased theoretical significance from reports that localized lesions of the brain can selectively impair either first or second order motion perception, while sparing the other, non-affected form (Vaina, Cowey, & Kennedy, 1999; Vaina, Grzywacz, LeMay, Bienfang, & Wolpow, 1998).

The correspondence problem

Some of motion perception's computational pre-requisites reflect computational obstacles to perception more generally. To appreciate this point, consider the problem of motion correspondence. It has long been recognized that the proximal stimulus under-specifies the distal stimulus. In describing this fact, Helmholtz noted that at any instant the distribution of light on the retina (the proximal stimulus) is consistent with an indefinitely large combination of stimulus objects and patterns of object illumination (distal stimuli). To resolve, or at least reduce, the proximal stimulus' massive under-specification of the distal stimulus, the visual system exploits various sources of supplementary information, including constraints and regularities embodied in the physics of the natural world.

Although motion perception must overcome the under-specification common to all perception, as a spatiotemporal process, motion faces some additional, unique sources of under-specification. An early stage of motion extraction requires correspondence matching. Some local pattern --for example, a luminance perturbation centered on retinal coordinates x_0, y_0 -- is detected at time t_0 and matched at later time, t_1 , to the pattern located at new coordinates, x_1, y_1 . As this description implies, motion depends upon a match or correspondence in space and time. As Attneave (1974) pointed out, when a stimulus comprises more than one element that moves, identification of correspondences over time becomes a real challenge.

Measurements of element positions in successive samples under-determine the samples' motion correspondences.¹⁰ Simple demonstrations of this point are shown in Figure 7. If there are n elements in each of two samples, then there are at least $n!$ sets of correspondences consistent with the samples. (This calculation assumes that one and only one element in the second sample is matched to each element in the first sample.) Dawson (1991) argued that to resolve this motion correspondence, the visual system exploits a trio of global spatiotemporal constraints, which mimic the properties of motion in the natural world. ("Global" implies that the constraints are applied

¹⁰ To illustrate the correspondence problem we have used examples in which stimuli are time-sampled at a relatively low rate. For historical reasons, when such stimuli generate the experience of motion, that motion is often designated as "apparent motion." It is important to note that despite our limited choice of illustrative examples, the correspondence problem confronts any biological motion system.

simultaneously to the entire field of dots, or to large portions of the field.) These constraints are known as the nearest neighbor principle (minimize the mean displacements between points matched between successive samples), the smoothness principle (because natural surfaces tend to vary smoothly, motion arising from the displacement of elements on such surfaces should be as smooth as possible (minimizing abrupt changes in velocity among neighboring elements), and the element integrity principle (because surfaces do not tend to pop in and out of existence, elements on such surfaces should persist over time --one element should not split into two, and two elements should not fuse into one) .

By emphasizing global motion matching, Dawson's model assumes that properties such as chromatic or shape similarities exert only a minor influence on motion correspondence. Although some empirical results have been taken as justifying that assumption, we believe that the preponderance of evidence leaves little doubt that motion correspondence is powerfully influenced by the similarity between stimuli. For example, A. B. Sekuler and Bennett (1996) examined motion correspondence generated by stimuli of differing, calibrated discriminability. Manipulating relative phase, a crucial determinant of form perception, they assessed strength of correspondence produced by quartets of stimuli. Each quartet comprised four stimuli evenly-spaced on a virtual circle. To emphasize the figural relationships among members of the quartet, we can designate the four stimuli as A, B, A, B.

Figure 8 illustrates that the entire quartet rotated rigidly over 45 degrees between successive display frames. Consider the four successive frames represented in that figure. If there were no color-based correspondence matching, the motion produced by the display would be ambiguous, as likely to move clockwise as counterclockwise. However, color-based matching would generate consistent motion in a clockwise direction. The elements in Sekuler and Bennett's quartets were compound sinusoidal gratings, each comprising a fundamental component and one harmonic component, at twice the frequency of the fundamental. Sekuler and Bennett manipulated the contrast and relative phase of the harmonic component in order to generate pairs of stimuli that varied in discriminability. When such stimuli were placed in quartet configuration (as in Figure 8), the probability of seeing motion in the direction dictated by figural matches varied with the perceptual difference between the two pairs of compound gratings. Tse, Cavanagh and Nakayama (1998) described a novel class of displays in which figural parsing and matching are requisites for perception of motion. ("Parsing" refers to the segmentation, spatial isolation and identification of separate stimulus elements.)

Before leaving the topic of motion correspondence, we should take special note of bistable motion sequences, that is stimuli in which the correspondence problem has two equally likely outcomes. An example of this unusual situation is illustrated in Figure 9, which shows two successive frames of animation sequence in which a pair of discs moves back and forth between two positions. Note that the upper disc in frame 1 could correspond either to the left-hand disc or to the right-hand disc in frame 2 (and likewise for the lower disc in frame 1). Because the two alternative motion paths are exactly equal in length, the motion system has no basis for deciding which pattern of motions is correct: motion correspondence is entirely ambiguous. When stimuli like this are viewed for an extended period, people perceive both possible paths (Anstis & Ramachandran, 1987; Kramer & Yantis, 1997; Pantle & Picciano, 1976; Ramachandran & Anstis, 1985), with each pattern of motion dominating for a few seconds, then giving way to the alternative pattern. This and other ambiguous motion displays are useful

tools for studying the stimulus factors that influence solutions to the correspondence problem (Francis & Kim, 1999; Yu, 2000). Multi-modal stimuli, such as auditory signals, can influence the perception of bistable motion displays (Lewkowicz, in press; Sekuler, Sekuler, & Lau, 1997; Watanabe & Shimojo, 2001).

The aperture problem

As described earlier, local motion is extracted from in the retinal image by neurons in Area V1. The receptive fields of such neurons can be construed as apertures, spatially delimited “windows” within which neurons register the presence of motion in a given direction. If an extended moving edge or bar is seen within such an aperture, no matter the direction the edge or bar actually moves, the neuron will respond as though the edge or bar were moving perpendicular to the orientation in the neuron’s receptive field (Bradley, 2001). To take a specific instance, imagine that a smooth, featureless, vertically-oriented bar moves obliquely, up and to the left at constant velocity. Any small receptive field positioned along the contour’s length signals only the motion component perpendicular to the contour’s orientation, in this case the leftward motion component ; because the contour extends beyond the boundaries of the small receptive field, the contour’s upward motion component produces no information that changes with time (Pack & Born, 2001).

Because each directionally-selective (DS) unit “sees” only what is happening within its own receptive field, the resulting signals from DS units are necessarily ambiguous. This ambiguity, now known as the aperture problem, was pointed out more than 60 years ago by Hans Wallach (Wuerger, Shapley, & Rubin, 1996). Figure 10 illustrates what Wallach had in mind. The circular area in the figure represents the receptive field of a directionally-selective neuron. Its preferred direction of movement is rightward. As the vertical edge of a large black line moves rightward at an appropriate speed through the receptive field, the neuron responds strongly (Panel A). But this isn’t the only direction of movement that could evoke such a response. So long as the black line is large compared to the aperture (the receptive field), the same local spatiotemporal event -- movement at the same velocity (meaning the same speed and direction) -- could be generated by any number of other combinations of direction and speed, some of which are suggested in Panels B and C. This equivalence means that the neuron’s response is inherently ambiguous. Because the neuron’s “view” of the world is limited to the confines of its receptive field, the neuron responds exactly the same way to each of these different velocities of movement. As a prelude to explaining the visual system’s strategy for resolving this ambiguity, consider one side effect of this ambiguity.

Movement seen within any single receptive field could have arisen from a variety of distinctly different visual events. This ambiguity opens the door to numerous illusions of perceived motion. With some of these, observers watch as a line or bar moves through an aperture whose shape strongly influences the perceived direction of movement. For example, suppose that a L-shaped aperture is cut in a piece of paper. Suppose also that a long oblique line (at 45 degrees) behind the aperture moves steadily straight downward (Figure 11A). Initially, an observer sees an oblique line that moves downward. Then, once the line reaches the base of the L-shaped aperture, the movement abruptly changes (Figure 11B). Now the observer sees a rightward moving oblique line. Wallach investigated this and several dozen, related aperture illusions (Wuerger, Shapley, & Rubin, 1996). To appreciate the basic elements of his approach,

return to the L-shaped aperture.

One can approximate the L-shaped aperture's upright portion by a set of co-linear circular receptive fields (Figure 11C). An oblique, downward-moving line would traverse each of these fields, one after another. But, this pattern of stimulation under-specifies the distal stimulus; in fact, the pattern of stimulation could have arisen from any number of distal stimuli. For example, it have been produced by several different, similarly-oriented lines, each of which traverses just one receptive field and then disappears, just as the next line appears at the top of the next receptive field and begins its own descent. Or, the same pattern of stimulation across all the receptive fields could have resulted, as it did in Wallach's demonstration, from a single line that moved from top to bottom, entering and then exiting one receptive field after another. How does the visual system, given the limited information available to it, select the scenario that most likely was responsible for that information? In the spirit of Gestalt psychology, Wallach proposed that the single perceptual choice made in such multiple-choice situations tends to be the simplest global motion. In this case, an observer sees a single, steadily-moving line rather than a series of different lines in succession. An alternative view might portray the single perceptual outcome as the product of a Bayesian perceptual inference (Knill, Kersten, & Yuille, 1996). Choosing between these alternative accounts requires additional, carefully-designed empirical measurements.

Perceived motion, then, is not determined solely by local responses to stimulus velocity generated within separate restricted regions of the field. Instead, local measurements of velocity are integrated at some place in the visual system at which the local velocity-related signals from area V1 are collected and combined. Such a combination of signals is the nervous system's standard operating procedure for resolving neural ambiguity in the signals of individual neurons. When the response of one neuron is ambiguous, the nervous system can diminish the ambiguity by aggregating outputs from a number of differently-tuned neurons. For visual motion, the nervous system reduces ambiguity about individual, local spatiotemporal events by channeling outputs from the first stage of direction processing to its second stage of neural processing, in MT.

Some neurons in MT receive input from V1 neurons with different preferred directions of motion. As a result, the directional-selectivity of these MT neurons differs qualitatively from that of their predecessors in V1. And these differences enhance the importance of MT's contribution to the perception of motion. To see this, return to Wallach's demonstration with two overlaid, moving gratings. When analogous two-component grating displays are imaged within the receptive fields of V1 neurons, the neurons tend to respond to the separate components. Some of Wallach's psychophysical observations foreshadowed these physiological results. He showed observers patterns composed of two different line gratings, like those shown Figure 12A and 12B. When the two are overlaid, as in Figure 12B they generate a series of diamond-shaped structures. When the bars of the two gratings move downward at the same rate, the display's appearance fluctuates between (a) a field of diamonds that moves downward (Figure 12C), and (b) two line-gratings that move at different velocities, sliding over each other in the process (Figure 12D). The relative potency of the two, alternative perceptual outcomes varies with a number of factors, Wallach found, including the angle between the two line gratings. This bistability is reminiscent of Rubin's vase/face illusion, or various phenomena of binocular rivalry (see Sekuler & Blake, 2001).

This perceptual bistability points to something quite important about motion

perception. In order to parse the display into two gratings moving in different directions, the nervous system requires mechanisms that separate out the different directions within each local region. But to produce the alternative outcome -- the moving diamond -- the nervous system needs a second, mechanism that collects dissimilar directional signals from regions across the display and binds those signals into coherent perceptual wholes rather than independent, isolated elements. This may sound challenging, but it is utterly necessary. In the example shown in Figure 12, these directions would be obliquely downward to the left and right. But when the same displays are presented in the receptive fields of MT neurons, many of those cells respond to the motion, not of the components, but of the overall moving pattern (Movshon, Adelson, Gizzi, & Newsome, 1986). In the case illustrated in Figure 12, this would be coherent motion directly downward. It is thought, then, that important aspects of motion perception arise from two stages of processing: one stage in which local motion vectors are extracted, and a second stage in which those vectors are sorted into object-related combinations. As Braddick put it “we do not live in a world of independent fragments, but of objects and surfaces that move coherently” (1993, p. 263). Vision needs some way to combine velocity signals that belong to the same moving object. It also needs to separate velocity signals that do not belong together, signals that arise from different objects (Snowden & Verstraten, 1999). Obviously, the characteristics of the world mandate precisely the sort of complex behavior that Wallach demonstrated.

Several computational strategies could be exploited in order to resolve the ambiguity represented by the aperture problem. Fennema and Thompson (1979) were first to observe that the ambiguity could be overcome by combining two or more ambiguous, local motion measurements. Movshon et al. (1986) demonstrated that some neurons in Area MT seem to perform precisely such a disambiguating computation, a computation producing what is known as an intersection of constraints (Adelson & Movshon, 1982).

Pack and Born (2001) demonstrated a dynamic form of disambiguation in macaque MT neurons. Initially, as expected from the influence of the aperture problem, MT neurons responded primarily to the component of motion perpendicular to the orientation of a moving contour. Then, over approximately 60 msec, the neurons' responses shifted steadily. At the end of this dynamic process, neurons' responses came to signal the actual direction of the stimulus, no matter what the stimulus' orientation. This shift in neural activity had a striking correlate in the monkeys' oculomotor behavior, in smooth pursuit eye movements. The initial velocity of pursuit eye movements deviates in a direction perpendicular to local contour orientation, suggesting that the earliest neural responses influence the oculomotor behavior. These results suggest that the primate visual system derives an initial estimate of motion direction by integrating ambiguous and unambiguous local motion signals over a fairly large spatial region, and refines this estimate over time. Lorenceau et al. (1993) demonstrated an analogous result with human observers: The perceived direction of a field of moving bars is initially perpendicular to the bars' orientation, but shifts slowly towards the actual direction of movement.

Geisler (1999) called attention to yet another avenue to the disambiguation of motion signals. With a stimulus that moves relatively rapidly, the the visual system's temporal integration produces an oriented virtual contour, which Geisler labelled a visual streak. (It is important to appreciate the virtual character of such streaks; they should not be confused with the actual streaks or “trails” sometimes generated as a

wake when a bright target moves rapidly across a cathode ray tube display.) The orientation selectivity of neurons in V1 means that motion streaks, which accurately reflect the direction of contour motion, could generate good, reliable orientation signals that can help disambiguate direction of motion (see also Jancke, 2000). Geisler buttressed his ingenious hypothesis with measurements of motion thresholds with simultaneous orientation masking and after orientation adaptation.

MOTION AFTEREFFECTS

When one stares at motion in one direction for a some period of time and then gazes at a stationary scene, the static scene appears to move in the direction opposite to the previously viewed motion. Although prolonged exposure to motion has other consequences --such as direction-selective threshold changes and alterations in perceived velocity-- illusory post-exposure motion defines what is known as the motion aftereffect (MAE).¹¹ Reports of this visual phenomenon can be traced back to Aristotle (about 330 B.C.) and Titus Lucretius Carus in about 56 B. C. (Wade & Verstraten, 1998). The best known early description of the phenomenon was given by Addams (1834), whose observation of the effect at Scotland's Falls of Foyer gave the phenomenon its common name, the waterfall illusion. Wade and Verstraten (1998) provide an excellent historical treatment of MAE, which is a widely used tool for studying motion processing more generally. In fact the explosive growth in the literature on MAE, as documented by Mather, Verstraten and Anstis (1998, p. viii) parallels the growth in published work on many aspects of visual motion.

The conditions under which MAE occurs were extensively studied by Wohlgemuth (1911). MAE's strength can be assessed by any of a number of measures, including apparent speed, apparent direction and duration (Pantle, 1998). Using MAE duration as an index of strength, Wolgemuth found that maintaining fixation during the adaptation period increased the effect and that adapting one eye would produce an aftereffect in the other, an example of interocular transfer. This surprising finding was initially reported by (Dvorak, 1870) and has been replicated recently by others (for example Heller & Ziefle, 1990; Raymond, 1993; Steiner, R, & Rose, 1994; Symons, Pearson, & Timney, 1996; Wade, Swanston, & De Weert, 1993). The interocular effects are usually only half as strong as the monocular effects.

Many motion studies generated adaptation with drifting sinusoidal grating, whose effect was measured with a stationary test grating. Not unexpectedly, the MAE was spatial frequency tuned: the strongest effect occurred when adapting the grating and the stationary test grating had the same spatial frequency (Bex, Verstraten, & Mareschal, 1996; Cameron, Baker, & Boulton, 1992; Thompson, 1998). The effect also showed temporal frequency tuning, with the strongest effect reported at a temporal frequency of 5 Hz, regardless of spatial frequency (Pantle, 1974).

While usually the aftereffect is evoked with a stationary test stimulus, more recent experiments have used dynamic random dot cinematograms as test stimuli, for

¹¹ Even very brief exposures to motion can generate substantial consequences, such as two-fold, directionally-selective decreases in sensitivity or clear alterations in perceived direction (Pantle, Gallogly, & Piehler, 2000).

example (Blake & Hiris, 1993; Ledgeway, 1994). Using such test stimuli, aftereffect strength can be measured by having a proportion of dots in the cinematogram move in the adapted direction to null the illusory motion. Moreover, adapting to non-Fourier motion does not produce an aftereffect with static tests (McCarthy, 1993; Nishida, Ashida, & Sato, 1994; von Grünau, 1986), but an after-effect is observed when the test stimulus is dynamic, such as a pattern of flickering elements (Ledgeway, 1994; McCarthy, 1993; Nishida & Sato, 1995). At least some investigators feel that dynamic, rather than static test patterns provide the most accurate reflection of motion mechanisms' properties (Hiris & Blake, 1992).

The MAE ordinarily lasts for a relatively brief time following adaptation, typically decaying after about 1- 15 seconds. Although the passage of time strongly influences the decay of the aftereffect, other factors also play a role. In fact, certain test viewing conditions can freeze the after-effect's normal decay, causing the after-effect to be stored (Spigel, 1960, 1962a, 1962b, 1964). Immediately after adaptation if one closes one's eyes for a period that exceeds the MAE's expected duration, when the eyes are reopened the MAE will still be seen. In other words, closing the eyes slows MAE's normal decay. This phenomenon is often described as an example of aftereffect "storage." Such storage has some interesting characteristics that have changed researchers' views of the MAE. For example, suppose that after adaptation, the MAE is tested with a dynamic field (randomly-moving dots). If the MAE is allowed to run its course with this dynamic test field, and then a static test field is presented, a MAE will be seen on that field. And, most surprising, the duration of the MAE on the static field will be little affected by the intervening, dynamic MAE. Reversing the order of the two test fields, with the static MAE followed by the dynamic MAE, the duration of the dynamic MAE is affected by the intervening static MAE (Verstraten, Fredericksen, van Wezel, Lankheet, & van de Grind, 1996). This relative independence of the static and dynamic MAEs, and other divergences between these two types of test field, has encouraged the idea that adaptation can occur in at least two distinct regions of the brain. Specifically, Nishida and colleagues (1994) suggested that the static MAE is due to adaptation in the primary motion detectors, while the dynamic MAE reflects adaptation in a higher cortical area such as MT.

Early modern accounts of the MAE assigned a central role to the fatiguing of cells during adaptation. These accounts exploited two key ideas: first, neurons in the motion system fire spontaneously in the absence of motion; and second, perceived direction of motion reflects the relative firing rates of neurons tuned to opposite directions. Suppose that the spontaneous activity of fatigued cells is significantly reduced for a brief time after adaptation. Therefore, the presentation of a stationary stimulus generates an imbalance in the spontaneous firing of the motion cells, with the non-adapted cells having a proportionately higher rate. According to this hypothesis, the brain takes this imbalance as signalling the presence of actual motion in a direction opposite to the earlier, adapting direction (Barlow & Hill, 1963; Sekuler & Ganz, 1963; Sutherland, 1961). And of course such illusory motion is what people perceive. This model was labelled the ratio model, because it held that perceived direction was controlled by the ratio of responses among motion cells tuned to different directions of motion. A somewhat-modified version of this theory, dubbed the distribution shift model (Mather, 1980), recognized that adaptation would affect a range of directionally-tuned mechanisms, not just the one mechanism that was most sensitive to the adapting direction. The result would be a more widespread change in post-adaptation spontaneous activity.

By incorporating various forms of inhibition, recent theories of the mechanisms of the MAE depart from earlier emphases on fatigue. In part, this new emphasis reflects a recognition of the central role that inhibition plays in the cortex, and, therefore, in visual computations. For example, Barlow (1990) proposed that the MAE results from a buildup of mutual inhibition among populations of motion-sensitive neurons during adaptation.¹² This mutual inhibition remains for some time after the adapting stimulus is removed. As a result, the pattern of responses among the motion cells is modified such that a neutral (stationary) stimulus is perceived as moving in the opposite direction. Since this inhibitory buildup could occur at any site where motion sensitive cells are found, adaptation can occur at one or more sites in the motion system. Nishida's MAE results (1995), mentioned earlier, are consistent with this idea.

Motion aftereffects can also be produced by stimuli comprising two superimposed directions of motion. Here, the resulting aftereffect is opposite the vector sum of the two adapting directions (Riggs & Day, 1980; Verstraten, Fredericksen, & van de Grind, 1994). Thus, for example, simultaneous adaptation to leftward motion and to upward motion subsequently causes a stationary test figure to appear to drift down and to the right (the vector sum of the downward component and the rightward component). When the two adapting components are unequal in strength --say, one component is higher contrast than the other-- the aftereffect direction shifts toward the stronger component, and the size of the shift tracks the inequality between stimulus components.

MEMORY AND LEARNING FOR MOTION

Vision's essential purpose is to guide behavior. And that purpose is served best when vision's current products can be compared to stored representations of vision's earlier products. Such comparisons, or recognitions, enable animals to prepare situation-appropriate behaviors in advance of their being needed. Although much of the memory research on humans has focused on memory for symbolic or verbal material, long before words came on the evolutionary stage, animals had a crucial need to remember and recognize scenes, objects and events they had encountered earlier. Although motion is certainly among the visual attributes that are worth remembering, relatively little research has been done on memory for motion. Magnussen and Greenlee (1992) examined observers' short-term memory for stimulus velocity. Using drifting luminance grating, they measured the difference threshold for velocity, V . They explored the retention of velocity information in memory by varying the interstimulus intervals (ISI) separating the first or reference stimulus, and the second or test stimulus. Weber fractions (V/V), proved to be independent of ISIs ranging from 1-30 seconds. This invariance showed that memory for velocity is quite robust over delays of up to 30 seconds.

Blake, Cepeda and Hiris (1997) explored memory for direction of movement, using random dot cinematograms with 100% coherence. After a one-second

¹² For Barlow (1990), the MAE exemplified a general principle of neural operation. In his view, the mutual inhibition, to which he ascribes MAE, reduces redundancy in the firing of different sets of neurons, thereby minimizing the total activity at any site in the visual system.

presentation of a cinematogram, an observer used a computer mouse to indicate the direction in which motion had been seen. The mean absolute value of observers' errors was about 5 degrees (Blake, Cepeda, & Hiris, 1997). This level of performance is remarkably robust when pre-response delays force the perceived direction to be stored and then retrieved from memory. For example, performance is unchanged when the subject's response is delayed by eight seconds. The memory on which responses were based is unlikely to be of an iconic nature; performance was unimpaired by random visual noise interposed between the cinematogram and the report of the remembered direction. Nor did memory depend on stored information from eye movement, such as observers might make while viewing the cinematogram: performance was unaffected when observers executed random tracking eye movements before making their judgments. Although memory for cinematogram direction was preserved over short intervals, such memory was dramatically undermined when Blake et al.'s observers saw not just one direction, but a succession of different directions on each trial. Performance fell off substantially as the number of seen directions grew. For example, average error climbed from 5 degrees with just one presentation, up to 25-30 degrees when observers saw and tried to remember seven different directions of motion.

Pasternak and colleagues have studied visual area MT's role in remembering the direction of visual motion (Bisley & Pasternak, 2000; Bisley, Zaksas, & Pasternak, 2001). They trained monkeys to compare the directions of motion portrayed in two successively presented animation sequences, the "sample" and the "test." Task difficulty was manipulated by varying the coherence of these random dot motion displays and the delay interval between sample and test. During the delay period -- while the monkey was trying to remember the sample direction that it had seen -- small, brief electrical pulses were delivered directly to clusters of neurons in MT. This electrical stimulation, which artificially activated the recipient neurons, influenced the monkey's subsequent judgment about the test direction (Bisley, Zaksas, & Pasternak, 2001). This result supports the notion that MT is involved in the short-term retention and/or retrieval of information about the direction of visual motion. This conclusion receives further support from the same researchers' studies of the effects of unilateral lesions to motion centers of the monkey cortex (Bisley & Pasternak, 2000).

As mentioned in this section's introduction, visual memory is an important guide for behavior. Memory of what we have seen allows us to prepare situation-appropriate behaviors and to execute them in a timely fashion. For visual movement, preparation of appropriate behavior requires recognition that some motion that is being experienced now has in fact been seen before. Chun and Jiang (1999) showed that, with repeated exposure to particular complex movement sequences promotes subsequent recognition of those sequences remains at the implicit level. In their experiment, observers searched for a single rotated-T target embedded in a field of L-shaped distractors. An animation sequence made all the search items, target as well as distractors, move randomly over the display screen; each item's random trajectory was independent of the trajectory of the other items. Because all the search items started the sequence as crosses, and only slowly morphed into their final form -- a rotated-T or L distractors -- observers had to monitor the entire set of randomly-moving items. Unbeknownst to observers, some of the random sequences were repeated during the 72 experimental trials. Search time improved also for all sequences, repeated as well as non-repeated, but the improvement was dramatically stronger with repeated sequences. Because explicit recognition accuracy was no better than chance, Chun and Yiang applied the label "implicit" to the perceptual learning they observed. This result

supports the notion that subjects can pick up and encode dynamic regularities or invariances, and that they can do so without explicitly recognizing those repeating features.

KNOWLEDGE, ATTENTION AND MOTION PERCEPTION

Attention takes many forms, all of which promote preferential processing of stimuli or stimulus attributes that are relevant to a particular task, and inhibit processing of task-irrelevant stimuli or stimulus attributes (Luck & Vecera, in press; Raymond, 2000). The selectivity that is represented by attention modulates the behavioral impact of any moving stimulus --up or down-- in accord with an observer's task and goals. Performance on any task, whether in laboratory or as part of everyday activity, implicitly reflects this modulation, which ordinarily operates in the background, unnoticed. In this section, we will consider a sample of experiments that were designed especially to highlight selectivity.

Tracking multiple objects.

William James distinguished among forms of attention, noting that while people can attend to sensory stimuli, such as particular locations or objects, they can also attend to "ideal or represented objects (1890); p. 393." In other words, James recognized that people could attend to a thing that was physically present, or they could attend to an object that existed only in the mind, what today might be called a "virtual object." Attention's ability to influence virtual moving objects is demonstrated by results from a procedure devised by Pylyshyn and Storm (1988). The procedure measures the ability to keep track, over time, of multiple, spatially dispersed, independently-moving targets. The task can be likened to watching some crows feeding in a field, and then, when the crows fly off in different directions, keeping track of each and every crow. In a typical experiment on multiple object tracking, a trial consists of three phases, called "target designation," "movement," and "probe." Figure 13 illustrates each of the three phases. Target designation began with the presentation of ten targets, such as discs, scattered about on a computer display. Then three, four or five of the ten discs, chosen randomly, blink several times. This informs the observer which discs will be targets for that trial. In the movement phase, all ten of the discs move about smoothly on the screen in various directions, each changing course unpredictably. After several seconds' movement, the trial's probe phase began. One of the ten discs is highlighted and the observer must report whether that one probe item had or had not been one of the targets designated, several seconds before, during the designation phase. Performance, summarized as the proportion of correct responses, is most accurate when people have to keep track of fewer targets, say three, rather than four or five. This is consistent with earlier findings that attention loses potency for any one item when additional items must also be attended to.

Yantis (1992) examined perceptual grouping's influence on multiple object tracking. For example, in one experiment, targets were either chosen at random or were chosen to lie at the vertices of virtual simple geometric figures, such as a diamond. At the outset of the experiment, performance was much superior when the target to be recognized had started out as part of a nice geometric perceptual group (during the target designation phase of a trial). However observers quickly learned to impose virtual groupings on elements that had not been part of ready-made, regular

geometrical groupings. This erased the early advantage found with ready-made groupings.

Unless grouping is maintained during the movement phase course, grouping at the start of a trial is no help at trial's end (the probe phase). By varying the targets' movements during the movement phase of a trial, Yantis controlled the likelihood that any existing grouping would be maintained. In one condition, targets moved about randomly, which allowed one or more targets to occasionally cross over an opposite edge of the virtual polygon. This criss-crossing destroyed the original grouping, undermining the coherence of the virtual polygon, and causing elements to lose identity as designated targets in that virtual polygon. In another condition, movements of targets were constrained, insuring that none ever crossed over an opposite edge of the virtual polygon. Here, movements of individual targets produced moment to moment fluctuations in the shape of a virtual figure that would be created by connecting those targets. But none of these fluctuations were drastic enough to destroy the convexity of the virtual polygon. Performance was distinctly better when the virtual polygon was preserved. This suggests that observers' attention creates (in the target designation phase) and maintains (in the movement phase) an updatable virtual object that is used (in the probe phase) to determine whether the probed target was or was not in the object.

To track multiple moving objects, the brain exploits neural circuits ordinarily dedicated to a different purpose, namely shifting attention from one location in space to another. Culham et al. (1998) used fMRI to identify the brain circuits that participated in multiple object tracking. Their results provide a clear picture of how the brain manages this difficult task. Attentive tracking of multiple, independently moving objects is mediated by a network of areas that includes parietal and frontal regions, known to be responsible for shifts of attention between locations and for eye movements, and area MT and related regions, which, as noted earlier, are central regions for processing motion information.

Uncertainty and motion detection

When it comes to detecting weak motion signals embedded in noise, it helps greatly to know in advance in what direction the signal dots are moving: uncertainty about direction of motion impairs detection performance. In one study, Ball and Sekuler (1981) determined the ease with which people detected very dim dots moving across a computer display. From trial to trial, the dots' direction of motion changed unpredictably. In addition, during half the trials, no dots at all were presented; the viewer saw just a blank screen. The dots were made dim enough that a viewer had great difficulty telling whether or not any dots were present. Ball and Sekuler measured the intensity threshold for detecting the dots under various conditions. Thresholds were initially determined simply by randomizing from trial to trial the direction in which the dots moved. Thresholds were then measured with an explicit cue that reduced the viewer's uncertainty about direction of motion. This directional cue was a short line flashed very briefly at different times relative to the presentation of the dots. The orientation of the line indicated the direction in which the dots, if present at all, might move (recall that on half the trials, no dots at all were presented).

Ball and Sekuler made several noteworthy discoveries. First, when the cue specified the dots' direction of motion precisely, the dots were easier to see -- that is, the intensity threshold was low. Second, the cue was not helpful unless it preceded the dots by about 500 msec, indicating that selective attention required some time to

operate. Third, if the cue's orientation did not match the dots' direction precisely, but only approximated it, the cue could still lower the detection threshold, but not as much as when it was precisely accurate. Generally, the greater the discrepancy between the cue's orientation and the dots' direction of motion, the more difficult it was to see the moving dots. In the extreme, a cue that misdirected the observer's expectation by 180 degrees, was worse than no cue at all: detection fell below a no-cue baseline. But how do directional cues or certainty about direction exert the effect that they do? Cues or prior knowledge are not part of the stimulus, but they certainly do affect the response evoked by the stimulus. Obviously, some information extracted from the cue must be recoded into a format capable of influencing subsequent processing of the test stimulus. After this recoding process, the non-moving cue then seems able to selectively boost responses in particular sets of directionally-selective neurons.

Shulman and colleagues (1999) extended Sekuler and Ball's study by measuring brain activity while people performed the cued-motion detection task. As before, the stationary cue, presented prior to the moving target, specified the direction of motion that people would see. Functional magnetic resonance imaging (fMRI) signals revealed that the non-moving cue activated brain areas that included MT, as well in adjacent regions that normally respond to motion. Also, some areas of parietal lobe, not normally responsive to motion, were activated. Together, these motion-sensitive and motion-insensitive areas constitute a neural circuit that encodes and maintains the cue during the interval between the cue and the onset of motion.

Presumably, prior information about the direction of motion temporarily boosts the signals of MT neurons that are particularly responsive to that direction of motion (Treue & Maunsell, 1999; Treue & Trujillo, 1999). This internally-generated boost response is equivalent to what happens when the responses of particular directionally-selective neurons are strengthened, either by the presentation of a strong visual stimulus or by direct electrical stimulation as in the study by Salzman et al. (1992). Sekuler, Sekuler and Sekuler (1990) used direction uncertainty to explore detection of changes in direction of movement. "Direction uncertainty" refers to an observer's prior knowledge of stimulus direction; as described above, uncertainty diminishes the detectability of motion, as indexed by elevated thresholds or lengthened reaction times. In this study, observers had to respond as quickly as possible to a constant relative change in stimulus direction --30 degree clockwise. The initial, pre-change direction was either fixed within a block of trials (certainty condition) or was completely random (maximum uncertainty). Generally, responses to change in the certainty condition were considerably faster than in conditions of uncertainty. However, if the initial, pre-change motion lasted 500 msec or longer, observers' reaction times to change were no longer affected by uncertainty about initial direction. However, for shorter initial durations, reaction time increased with increased uncertainty (i. e., increases in the range of possible initial directions). Sekuler et al. proposed that the visual system requires approximately 500 msec to normalize the initial direction of motion to be able to detect the direction change by essentially converting the nominal task to one of detecting motion onset.

Alais and Blake (1999) used the MAE to probe attention's influence on motion perception. As mentioned earlier, when the adapting stimulus comprises two directional components, the direction of the MAE is usually the vector sum of the two components, but when the components are unequal in strength, the resulting MAE tracks that inequality. By varying the relative contrasts of the two component adapting motions, Alais and Blake were able to manipulate how much attention an observer had

to pay to one of the components. In their experiments, observers viewed a computer display consisting of two superimposed fields of moving dots. In one group, all dots moved coherently in a single direction, shown as “upward” in Figure 14. Dots in the other group moved in random directions most of the time, producing no net directional drift. Every once in a while, a subset of the random dots joined forces to move in the same direction, shown as “rightward” in Figure 14. The proportion of dots moving rightward was only about 25%, making it necessary for observers to look carefully to detect their presence. On some trials, observers were required to indicate when this weak coherent motion was present; on the remaining trials observers simply observed the same stimuli but did not have to perform the detection task and, therefore, did not have to attend selectively to the rightward motion. Alais and Blake reasoned that the first of these conditions would demand more attention than the second, passive viewing condition. In the passive, control condition, the brief insertions of weak rightward motion had little effect on MAE’s direction; however, when observers had to attend to it, the inserted motion dramatically altered the aftereffects’ direction, shifting it by about 20 degrees. This same shift, which was mediated solely by observers’ attention to a weak signal, was equivalent to the shift that would have been produced by a very powerful signal (motion with dots 70% correlated). So attention to motion in one direction boosted the response to that motion by almost threefold --rendering a 25% correlated stimulus as effective as one that was correlated 70%. Extrapolating from neural responses within MT to changes in degree of stimulus correlation, Alais and Blake deduced that the attentional effects seen in human observers were equivalent, on average, to what would be expected from doubling the stimulus correlation in a non-attended stimulus. There’s no doubt that attention can exert a very powerful influence on perceived motion.

Parsing everyday activities. Although considerable progress has been made toward an understanding of motion perception, a number of important questions remain unresolved. Among them is the question of how human observers recognize and categorize everyday actions. To clarify the problem, Bobick (1997) has distinguished between what he calls movements, such as the lifting of arm, and what he calls actions, which include interactions with the environment and other actors, as well as inferred causal relationships among image components. To take just a few examples, actions include a soccer player’s heading the ball, a cook’s making a cheese sandwich, and your making a bed or ironing a shirt. In all these cases, the actor (the soccer player, the cook, and you) generates a more-or-less continuous stream of movements. An observer’s understanding of the visual information could begin with a parsing of the action into distinct perceptual events (Tong, 2001). One perceptual component event could be distinguished from a succeeding component by changes in velocity and/or movement of one body part relative to another. So motion perception is a possible for our ability to parse complex actions. Combining behavioral observations and a novel functional brain imaging paradigm, Zacks and colleagues (2001) examined motion’s possible contribution to action parsing. Participants first watched movies of everyday activities. The movies, 2-5 minutes long, showed someone making a bed, fertilizing a houseplant, ironing a shirt, or washing dishes. Movies were viewed three times each, in random order. During the first viewing of any movie, observers merely watched the action passively. During this viewing, fMRI measurements were taken. Then, during a movie’s subsequent viewings, participants used a button to signal when they thought one that one natural and meaningful unit of action had ended and another one had

begun. During the original, passive viewing of everyday activities, fMRI signals reflected transient changes occurring in several, related regions of the brain, including Area MT+, which participates in motion perception. The onset of transient changes in neural activity did not occur randomly during the action, but were in temporal register with moments that observers deemed to be boundaries between components of the overall action. So it may be that motion information plays a key role in the segmentation and understanding of everyday actions.

It is worth noting when these transient changes in MT+ activation occurred. On average, they began a few seconds before the perceived boundary between action components. As a result, it could be that these anticipatory transient changes in brain activation signify top-down influences, that is influences of observers' familiarity with the actions and therefore observers' expectancies about upcoming changes in motion. Support for this hypothesis comes from fMRI research with motion that is imagined, but not actually seen (Grossman & Blake, 2001; Kourtzi & Kanwisher, 2000).

Summing up and looking forward

Without question, our understanding of motion perception has been pushed to a level scarcely imaginable just two decades ago. New psychophysical, physiological, and computational research tools, have made possible huge strides toward unraveling the mysteries of the visual registration of motion, which Walls (1942) and we consider to be most important among the talents that comprise vision. The application of functional neuroimaging has begun to identify rough, but intriguing, relationships between particular sites in the visual brain and performance of particular tasks. But obviously, this development has only just started, and many difficult challenges lie just ahead.

Our present knowledge of links between brain sites and aspects of visual motion perception throws only the dimmest of lights onto the complex neural transformations and computations that support performance on various tasks. Motion perception emerges from an shifting partnership between exogenous influences, represented by stimulus attributes, and endogenous influences, including expectation, attention, memory, and learning. We have relatively little understanding of this partnership's parameters and its limiting conditions. Also, we have no understanding whatever of the control architecture that sets and adjusts the relative weights for the two partners, exogenous and endogenous influences. And it is not clear to us that traditional behavioral paradigms will have much to contribute to the development of such understanding.

Vernon Mountcastle (quoted in Shadlen & Newsome, 1996) sketched out an ambitious, broad agenda not only for the study of visual motion, but also for the entire field of sensory science. He urged researchers to study the complete chain of events that "lead from the initial central representation of sensory stimuli, through the many sequential and parallel transformations of those neural images, to the detection and discrimination processes themselves, and to the formation of general commands for behavioral responses and detailed instructions for their motor execution." This last part of Mountcastle's agenda would connect decisions about sensory signals to the preparation and execution of motor acts appropriate to those decisions, an exciting area in which work has only just begun (see, for example, Gold & Shadlen, 2001).

Figure Legends

Figure 1. Four still frames cut from a video by Bobick and Davis (2001). The video shows a person engaged in a common, everyday activity. The low pass spatial filtering of video makes it difficult, from any individual frame, to discern the person, let alone what the person is doing. However, when the video is played at normal rate, the pattern of motions make both the person and the person's action immediately apparent.

Figure 2. Diagram illustrating proposed functional and anatomical streams in the primate cerebral cortex. Partially separate streams carry information from area V1 either dorsally, toward the inferior parietal cortex (PG), or ventrally, toward the inferior temporal cortex (TE). Arrows indicate the main feedforward projections between areas. Abbreviations used in the diagram: V1, primary or striate cortex; MT, middle temporal area (also known as V5); VIP, ventral intraparietal; LIP, lateral intraparietal; PP, posterior parietal, MST, medial superior temporal; FST, fundus superior temporal; PG, inferior parietal cortex; TE, inferior temporal cortex. (After Ungerleider & Haxby, 1994).

Figure 3. Illustrative space-time (x-t) diagrams that are used to represent motion. Panel A. Vertical bar moves rightward at constant speed. Panel B. The space-time representation of the movement in Panel A. Panel C. Space-time representation for bar moving rightward at higher speed than in Panel A. Panel D. Space-time representation for bar moving leftward at same speed as in Panel A. Panel E. Space-time representation for bar that moves rightward, suddenly stops and remains stationary. Panel F. Space-time representation for bar that moves rightward and then abruptly reverses direction.

Figure 4. Diagram of display used by Watamaniuk et al (1995) to examine effects of occlusion on trajectory detection. See text.

Figure 5. Panel A. Optic flow pattern produced when an observer translates along a straight path; diagram assumes that observer's is directed toward destination point. Panel B. Optic flow pattern produces when an observer translates along a curved path. (From Warren, Blackwell, & Morris, 1988).

Figure 6. Two-dimensional projection (right) of transparent rotating cylinder (left). The dots on the cylinder's surface create intermingled left and right moving dots in the two-dimensional projection.

Figure 7. Illustration of motion's correspondence problem. In each panel, the three black squares represent items presented in the first frame of a cinematogram, the three white squares represent items in the second frame. Panels illustrate various ways that items from the two frames could be matched. (After Dawson, 1991)

Figure 8. Four frames of a quartet display. In each frame, two tokens are black and two are gray; the quartet of tokens rotates rigidly by 45 degrees between successive frames. In the absence of color-based correspondence matching, the display's motion would be ambiguous, as likely to move clockwise as counterclockwise. However, color-based matching generates consistent motion in a clockwise direction, with the probability of seeing motion in the direction dictated by color matches increasing with the perceptual difference between the two pairs of tokens.

Figure 9. Two successive frames of a display that illustrates feature-matching in solution of correspondence problem. The upper disc in frame 1 could correspond either to frame 2's left-hand disc or to its right-hand disc (and likewise for the

lower disc in frame 1). Because the two alternative motion paths are equal in length, the motion system has no basis for deciding which pattern of motions is correct: motion correspondence is entirely ambiguous, were it not for feature (color) correspondence. When such stimuli are viewed for an extended period, both possible paths are seen in alternation.

Figure 10. Illustration of the aperture problem: The ambiguity of directional information within a receptive field aperture. The circular area represents the receptive field of a directionally-selective neuron tuned to rightward motion. As the vertical bar moves rightward through the receptive field, the neuron signals rightward motion (Panel A). However, movement of the same bar downward and to the right (Panel B) also generates rightward movement within the receptive field, and could movement upward and to the right (Panel C). The equivalence of these and other directions of actual movement renders the neuron's response ambiguous. (Modified from Sekuler & Blake, 2001)

Figure 11. Wallach's demonstration of aperture shape's influence on perceived motion. Panel A: Motion an obliquely oriented line into, through, and then out of an L-shaped aperture. Panel B suggests that the line initially appears to move vertically downward; then, when the line enters the aperture's horizontally-oriented portion, its apparent direction abruptly changes to horizontal. See text for explanation of Panel C. (Modified from Sekuler & Blake, 2001)

Figure 12. Superimposition of two diagonal, moving gratings (A and B) produces a diamond pattern whose motion is bistable. When either moving grating is presented alone, it appears to move in a consistent direction, which is indicated by the arrows in A and B. When the gratings are superimposed, sometimes they cohere to form a single, downward moving pattern (Panel C) and, at other times, they appear to slide over one another, in two different superimposed directions. (Demonstration devised by Wallach. Diagram modified from Sekuler & Blake, 2001)

Figure 13. Typical trial structure for multiple target tracking experiment. During trial's initial phase (target designation), several discs are singled out by a momentary brightening. During the second phase (movement), all discs undergo a series of random movements. In phase 3 (probe), a single disc is brightened, and the observer must judge whether that disc was or was not among the ones that had been identified as potential targets during the trials's initial phase. (Modified from Sekuler & Blake, 2001)

Figure 14. Schematic of stimulus conditions used by Alais and Blake (1995) to test attention's influence on motion adaptation. During most of the adaptation period, a subset of adaptation dots moved coherently upward (open circles in left panel), while the remaining, noise dots moved randomly in all directions (gray circles in left panel). Unpredictably, a subset of the noise dots moved briefly rightward (black dots in center panel), and then quickly resumed their normal motions (right panel). Observers had to detect this brief motion, which required considerable attention. Note that throughout the experiment all dots were actually the same color; color differences are used here for illustrative purposes only. (Modified from Sekuler & Blake, 2001)

- Addams, R. (1834). An account of a peculiar optical phenomenon seen after having looked at a moving body. London and Edinburgh Philosophical Magazine and Journal of Science, *5*, 373-374.
- Adelson, E. H., & Movshon, J. A. (1982). Phenomenal coherence of moving visual patterns. Nature, *300*, 523-525.
- Ahlström, U., & Börjesson, E. (1996). Segregation of motion structure from random noise. Perception, *25*, 279-291.
- Ahlström, V., Blake, R., & Ahlström, U. (1997). Perception of biological motion. Perception, *26*, 1539-1548.
- Alais, D., & Blake, R. (1999). Neural strength of visual attention gauged by motion adaptation. Nature Neuroscience, *2*, 1015-1018.
- Albensi, B. C., & Powell, J. H. (1998). The differential optomotor response of the four-eyed fish *Anableps anableps*. Perception, *27*, 1475-1483.
- Andersen, G. J. (1989). Perception of three-dimensional structure from optic flow without locally smooth velocity. Journal of Experimental Psychology: Human Perception & Performance, *15*, 363-371.
- Andersen, R. A., & Bradley, D. C. (1998). Perception of three-dimensional structure from motion. Trends in Cognitive Science, *2*, 222-228.
- Anstis, S., & Ramachandran, V. S. (1987). Visual inertia in apparent motion. Vision Research, *27*, 755-764.
- Arterberry, M. E., & Yonas, A. (2000). Perception of three-dimensional shape specified by optic flow by 8-week-old infants. Perception & Psychophysics, *62*, 550-556.
- Attneave, F. (1974). Apparent movement and the what-where connection. Psychologia, *17*, 108-120.
- Baddeley, R., & Tripathy, S. P. (1998). Insights into motion perception by observer modeling. Journal of the Optical Society of America A, *15*, 289-296.
- Ball, K., & Sekuler, R. (1981). Cues reduce direction uncertainty and enhance motion detection. Perception & Psychophysics, *30*, 119-128.
- Ball, K., & Sekuler, R. (1986). Improving visual perception in older observers. Journal of Gerontology, *41*, 176-182.
- Ball, K., Sekuler, R., & Machamer, J. (1983). Detection and identification of moving targets. Vision Research, *23*, 229-238.
- Barlow, H. B. (1978). The efficiency of detecting changes of density in random dot patterns. Vision Research, *18*, 637-650.
- Barlow, H. B. (1980). The absolute efficiency of perceptual decisions. Philosophical Transactions of the Royal Society of London (Series B), *290*.
- Barlow, H. B. (1990). A theory about the functional role and synaptic mechanism of visual after effects. In C. Blakemore (Ed.), Coding and Efficiency (pp. 363-375). Cambridge, UK: Cambridge University Press.
- Barlow, H. B. (1995). The neuron doctrine in perception. In M. S. Gazzaniga (Ed.), The Cognitive Neurosciences (pp. 415-435). Cambridge: MIT Press.
- Barlow, H. B., & Hill, R. M. (1963). Evidence for a physiological explanation for the waterfall phenomenon and figural aftereffects. Nature, *200*, 1345-1347.
- Barlow, H. B., & Reeves, B. C. (1979). The versatility and absolute efficiency of detecting mirror symmetry in random dot displays. Vision Research, *19*, 783-793.

- Barlow, H. B., & Tripathy, S. P. (1997). Correspondence noise and signal pooling in the detection of coherent visual motion. Journal of Neuroscience, *17*, 7954-7966.
- Beckers, G., & Homberg, V. (1992). Cerebral visual motion blindness: transitory akinetopsia induced by transcranial magnetic stimulation of human area V5. Proceedings of the Royal Society of London, Series B, *249*, 173-178.
- Bennett, P. J., Sekuler, A. B., & Ozin, L. (1999). Effects of aging on calculation efficiency and equivalent noise. Journal of the Optical Society of America A, *16*, 654-668.
- Bex, P. J., Verstraten, F. A. J., & Mareschal, I. (1996). Temporal and spatial frequency tuning of the flicker motion aftereffect. Vision Research, *36*, 2721-2727.
- Bichot, N. P., Thompson, K. G., Chenthal Rao, S., & Schall, J. D. (2001). Reliability of macaque frontal eye field neurons signaling saccade targets during visual search. Journal of Neuroscience, *21*, 713-725.
- Bischof, W. F., Reid, S. L., Wylie, D. R., & Spetch, M. L. (1999). Perception of coherent motion in random dot displays by pigeons and humans. Perception & Psychophysics, *61*, 1089-1101.
- Bisley, J. W., & Pasternak, T. (2000). The multiple roles of visual cortical areas MT/MST in remembering the direction of visual motion. Cerebral Cortex, *10*, 1053-1065.
- Bisley, J. W., Zaksas, D., & Pasternak, T. (2001). Microstimulation of cortical area MT affects performance on a visual working memory task. Journal of Neurophysiology, *85*, 187-196.
- Blake, R. (1993). Cats perceive biological motion. Psychological Science, *4*, 54-57.
- Blake, R., & Aiba, T. S. (1998). Detection and discrimination of optic flow components. Japanese Psychological Research, *40*, 19-30.
- Blake, R., Cepeda, N. J., & Hiris, E. (1997). Memory for visual motion. Journal of Experimental Psychology: Human Perception & Performance, *23*, 353-369.
- Blake, R., & Hiris, E. (1993). Another means for measuring the motion aftereffect. Vision Research, *33*, 1589-1592.
- Bobick, A., & Davis, J. (2001). The recognition of human movement using temporal templates. IEEE Transactions on Pattern Analysis and Machine Intelligence, *23*, 257-268.
- Bobick, A. F. (1997). Movement, activity and action: the role of knowledge in the perception of motion. Philosophical Transactions of the Royal Society of London (Series B), *352*, 1257-1265.
- Borst, A. (2000). Models of motion detection. Nature Neuroscience (Supplement), *3*, 1168.
- Braddick, O. (1993). Segmentation versus integration in visual motion processing. Trends in Neuroscience, *16*, 263-268.
- Bradley, D. (2001). MT signals: Better with time. Nature Neuroscience, *4*, 346-348.
- Bravo, M. J., & Watamaniuk, S. N. J. (1995). Evidence for two speed signals: a coarse local signal for segregation and a precise global signal for discrimination. Vision Research, *35*, 1691-1697.
- Britten, K. H., & van Wezel, R. J. (1998). Electrical microstimulation of cortical area MST biases heading perception in monkeys. Nature Neuroscience, *1*, 59-63.
- Brown, R. H. (1961). Visual sensitivity to differences in velocity. Psychological Bulletin, *58*, 89-103.
- Brownlow, S., Dixon, A. R., Egbert, C. A., & Radcliffe, R. D. (1997). Perception of movement and dancer characteristics from point-light displays of dance. Psychological Record, *47*, 411-421.

- Buracas, G. T., Zador, A. M., DeWeese, M. R., & Albright, T. D. (1998). Efficient discrimination of temporal patterns by motion-sensitive neurons in primate visual cortex. Neuron, *20*, 959-969.
- Burgess, A. E., & Barlow, H. B. (1983). The efficiency of numerosity discrimination in random dot images. Vision Research, *23*, 811-829.
- Burgess, A. E., Wagner, R. F., Jennings, R. J., & Barlow, H. B. (1981). Efficiency of human visual signal discrimination. Science, *214*, 93-94.
- Burr, D. C., Morrone, M. C., & Vaina, L. M. (1998). Large receptive fields for optic flow detection in humans. Vision Research, *38*, 1731-1743.
- Cameron, E. L., Baker, C. L., & Boulton, J. C. (1992). Spatial frequency selective mechanisms underlying the motion aftereffect. Vision Research, *32*, 561-568.
- Celebrini, S., & Newsome, W. T. (1994). Neuronal and psychophysical sensitivity to motion signals in extrastriate area MST of the macaque monkey. Journal of Neuroscience, *14*.
- Chubb, C., & Sperling, G. (1988). Drift-balanced random stimuli: A general basis for studying non-Fourier motion perception. Journal of the Optical Society of America A, *5*, 1986-2006.
- Chun, M. M., & Jiang, Y. (1999). Top-down attentional guidance based on implicit learning of visual covariation. Psychological Science, *10*, 360-365.
- Cowey, A., & Vaina, L. M. (2000). Blindness to form from motion despite intact static form perception and motion detection. Neuropsychologia, *38*, 566-578.
- Croner, L. J., & Albright, T. D. (1997). Image segmentation enhances discrimination of motion in visual noise. Vision Research, *37*, 1415-1427.
- Croner, L. J., & Albright, T. D. (1999). Seeing the big picture: integration of image cues in the primate visual system. Neuron, *24*, 777-789.
- Culham, J., He, S., Dukelow, S., & Verstraten, F. A. J. (2001). Visual motion and the human brain: What has neuroimaging told us? Acta Psychologica, *107*, 69-94.
- Culham, J. C., Brandt, S. A., Cavanagh, P., Kanwisher, N. G., Dale, A. M., & Tootell, R. B. (1998). Cortical fMRI activation produced by attentive tracking of moving targets. Journal of Neurophysiology, *80*, 2657-2670.
- Curran, W., & Braddick, O. J. (2000). Speed and direction of locally-paired dot patterns. Vision Research, *40*, 2115-2124.
- Cutting, J. E. (1986). Perception with an eye for motion. Cambridge: MIT Press.
- Dakin, S. C., & Mareschal, I. (2000). The role of relative motion computation in 'direction repulsion'. Vision Research, *40*, 833-841.
- Dawson, M. R. W. (1991). The how and why of what went where in apparent motion: Modeling solutions to the motion correspondence problem. Psychological Review, *98*, 569-603.
- De Bruyn, B., & Orban, G. A. (1993). Segregation of spatially superimposed optic flow components. Journal of Experimental Psychology: Human Perception and Performance, *19*, 1014-1027.
- DeAngelis, G., Cumming, B. G., & Newsome, W. T. (1998). Cortical area MT and the perception of stereoscopic depth. Nature, *394*, 677-680.
- DeBruyn, B., & Orban, G. A. (1988). Human velocity and direction discrimination measured with random dot patterns. Vision Research, *28*, 1323-1335.
- Derrington, A. M., & Henning, G. B. (1993). Detecting and discriminating the direction of motion of luminance and colour gratings. Vision Research, *33*, 799-811.
- Diener, H. C., Wist, E. R., Dichgans, J., & Brandt, T. H. (1976). The spatial frequency effect on perceived velocity. Vision Research, *16*, 169-176.

- Dobkins, K. R. (2000). Moving colors in the lime light. Neuron, *25*, 15-18.
- Dror, R. O., O'Carroll, D. C., & Laughlin, S. B. (2001). Accuracy of velocity estimation by Reichardt correlators. Journal of the Optical Society of America A, *182*, 241-252.
- Dvorak, V. (1870). Versuche über Nachbilder von Reizveränderungen. Sitzungsberichte der Wiener Akademie der Wissenschaften, *61*, 257-262.
- Dzhafarov, E. N., Sekuler, R., & Allik, J. (1993). Detection of changes in speed and direction of motion: reaction time analysis. Perception & Psychophysics, *54*, 733-750.
- Emerson, R. C., Bergen, J. R., & Adelson, E. H. (1992). Directionally selective complex cells and the computation of motion energy in cat visual cortex. vision Research, *32*, 203-218.
- Exner, S. (1888). Über optische Bewegungsempfindungen. Biologisches Centralblatt, *8*, 437-448.
- Fennema, C. I., & Thompson, W. B. (1979). Velocity determination via scenes containing several moving images. Computer Graphics and Image Processing, *9*, 301-315.
- Field, D. J., Hayes, A., & Hess, R. F. (1993). Contour integration by the human visual system: evidence for a local "association field". Vision Research, *33*, 173-193.
- Fox, R., & McDaniel, C. (1982). The perception of biological motion by human infants. Science, *218*, 486-487.
- Francis, G., & Kim, H. (1999). Motion parallel to line orientation: disambiguation of motion percepts. Perception, *28*, 1243-1255.
- Gabbiani, F., Mo, C., & Laurent, G. (2001). Invariance of angular threshold computation in a wide-field looming-sensitive neuron. Journal of Neuroscience, *21*, 314-329.
- Gegenfurtner, K. R., & Hawken, M. J. (1996). Interaction of motion and color in the visual pathways. Trends in Neuroscience, *19*, 394-401.
- Gegenfurtner, K. R., Mayser, H. M., & Sharpe, L. T. (2000). Motion perception at scotopic light levels. Journal of the Optical Society of America A, *17*, 1505-1515.
- Geisler, W. S. (1989). Sequential ideal-observer analysis of visual discriminations. Psychological Review, *96*, 267-314.
- Geisler, W. S. (1999). Motion streaks provide a spatial code for motion direction. Nature, *400*, 65-69.
- Geisler, W. S., Perry, J. S., Super, B. J., & Gallogly, D. P. (2001). Edge co-occurrence in natural images predicts contour grouping performance. Vision Research, *41*.
- Gibson, E. J., Gibson, J. J., Smith, O. W., & Flock, H. (1959). Motion parallax as a determinant of perceived depth. Journal of Experimental Psychology, *58*, 40-51.
- Gold, J., Bennett, P. J., & Sekuler, A. B. (1999). Identification of band-pass filtered letters and faces by human and ideal observers. Vision Research, *39*, 3537-3560.
- Gold, J. I., & Shadlen, M. N. (2001). Neural computations that underlie decisions about sensory stimuli. Trends in Cognitive Science, *5*, 10-16.
- Graham, C. H. (1951). Perception of motion. In S. S. Stevens (Ed.), The Handbook of Experimental Psychology. New York: J Wiley.
- Graham, C. H. (1965). Perception of motion. In C. H. Graham (Ed.), Vison and Visual Perception (pp. 575-588). New York: John Wiley & Sons.
- Graham, N. V. S. G. (1989). Visual pattern analyzers. New York: Oxford University Press.

Gros, B. L., Blake, R., & Hiris, E. (1998). Anisotropies in visual motion perception: a fresh look. Journal of the Optical Society of America A, *15*, 2003-2011.

Grossman, E. D., & Blake, R. (1999). Perception of coherent motion, biological motion and form-from-motion under dim-light conditions. Vision Research, *39*, 3721-3727.

Grossman, E. D., & Blake, R. (2001). Brain activity evoked by inverted and imagined biological motion. Vision Research, *41*, 1475-1482.

Grossman, E. D., Donnelly, M., Price, P., Morgan, V., Pickens, D., Neighbor, G., & Blake, R. (2000). Brain areas involved in perception of biological motion. Journal of cognitive Neuroscience, *12*, 711-720.

Grzywacz, N. M., Watamaniuk, S. N. J., & McKee, S. P. (1995). Temporal coherence theory for the detection and measurement of visual motion. Vision Research, *35*, 3183-3203.

Hecht, S., Shlaer, S., & Pirenne, M. H. (1942). Energy, quanta, and vision. Journal of General Physiology, *25*, 819-840.

Heeger, D. J. (1994). The representation of visual stimuli in primary visual cortex. Current Directions in Psychological Science, *3*, 159-163.

Heeger, D. J., Boynton, G. M., Demb, J. B., Seidemann, E., & Newsome, W. T. (1999). Motion opponency in visual cortex. Journal of Neuroscience, *19*, 7162-7174.

Heeger, D. J., Simoncelli, E. P., & Movshon, J. A. (1996). Computational models of cortical visual processing. Proceeding of the National Academy of Science USA, *93*, 623-627.

Heidenreich, S. M. Z., G L (1995) Perception, *24*, 879-890 . (1995). Evidence that luminant and equiluminant motion signals are integrated by directionally selective mechanisms. Perception, *24*, 879-890.

Heller, D., & Ziefle, M. (1990). Zum interokularen Transfer beim Bewegungsnacheffekt. Würzburg, Germany.

Hikosaka, O., Miyauchi, S., & Shimojo, S. (1993). Visual attention revealed by an illusion of motion. Neuroscience Research, *18*, 11-18.

Hildreth, E. C. (2001). Personal communication to R Sekuler, May 2001 Boston.

Hiris, E., & Blake, R. (1992). Another perspective on the visual motion aftereffect. Proceeding of the National Academy of Science USA, *89*, 9025-9028.

Hiris, E., & Blake, R. (1996). Direction repulsion in motion transparency. Visual Neuroscience, *13*, 187-197.

Hochberg, J. (1988). Visual perception. In R. C. Atkinson & R. J. Herrnstein & G. Lindzey & R. D. Luce (Eds.), Stevens' Handbook of Experimental Psychology (Second ed., Vol. 1, pp. 195-276). New York: Wiley Interscience.

Hol, K., & Treue, S. (2001). Different populations of neurons contribute to the detection and discrimination of visual motion. Vision Research, *41*, 685-689.

Hotson, J. R., & Anand, S. (1999). The selectivity and timing of motion processing in human temporo-parieto-occipital and occipital cortex: a transcranial magnetic stimulation study. Neuropsychologia, *37*, 169-179.

Hubel, D. H., & Wiesel, T. N. (1968). Receptive fields and functional architecture of monkey striate cortex. Journal of Physiology, *195*, 215-243.

James, W. (1890). Principles of Psychology. New York: H. Holt.

Jancke, D. (2000). Orientation formed by a spot's trajectory: A two-dimensional population approach in primary visual cortex. Journal of Neuroscience, *20*, RC86(81-86).

Johansson, G. (1973). Visual perception of biological motion and a model for its analysis. Perception & Psychophysics, *14*, 201-211.

- Johnston, A., & Wright, M. J. (1986). Matching velocity in central and peripheral vision. Vision Research, *26*, 1099-1109.
- Katz, E., Gizzi, M. S., Cohen, B., & Malach, R. (1990). The perceived speed of an object is affected by the distance it travels. Perception, *19*, 387.
- Kirschen, M. P., Kahana, M. J., Sekuler, R., & Burack, B. (2000). Optic flow helps humans learn to navigate through synthetic environments. Perception, *29*, 801-818.
- Knill, D. C., Kersten, D., & Yuille, A. (1996). A Bayesian formulation of visual perception. In D. C. Knill & W. Richards (Eds.), Perception as Bayesian Inference.: Cambridge University Press.
- Koenderink, J. J. (1986). Optic flow. Vision Research, *26*, 161-180.
- Kourtzi, Z., & Kanwisher, N. G. (2000). Activation in human MT/MST by static images with implied motion. Journal of cognitive Neuroscience, *12*, 48-55.
- Kozlowski, L. T., & Cutting, J. E. (1977). Recognizing the gender of walkers from dynamic point-light displays. Perception & Psychophysics, *21*, 575-580.
- Kramer, P., & Yantis, S. (1997). Perceptual grouping in space and time: evidence from the Ternus display. Perception & Psychophysics, *59*, 87-99.
- Krekelberg, B., & M, L. (2001). Neuronal latencies and the position of moving objects. Trends in Neuroscience, *24*, 335-339.
- Lappin, J. S., Donnelly, M. P., & Kojima, H. (2001). Coherence of early motion signals. Vision Research, *41*, 1631-1644.
- Ledgeway, T. (1994). Adaptation to second-order motion results in a motion aftereffect for directionally-ambiguous test stimuli. Vision Research, *34*, 2879-2889.
- Lee, D. N. (1976). A theory of visual control of braking based on information about time-to-collision. Perception, *5*, 437-459.
- Lee, D. N. (1980). The optic flow field: The foundation of vision. Philosophical Transactions of the Royal Society of London (Series B), *290*, 169-179.
- Lee, S. H., & Blake, R. (1999). Detection of temporal structure depends on spatial structure. Vision Research, *39*, 3033-3048.
- Leibowitz, H. (1955). Effect of reference lines on the discrimination of movement. Journal of the Optical Society of America, *45*, 829-830.
- Levinson, E., & Sekuler, R. (1975). The independence of channels in human vision selective for direction of movement. Journal of Physiology, *250*, 347-366.
- Lewkowicz, D. J. (in press). Heterogeneity and heterochrony in the development of intersensory perception. Cognitive Brain Research.
- Lindsey, D., & Todd, J. T. (1998). Opponent motion interactions in the perception of transparent motion. Perception & Psychophysics, *60*, 558-574.
- Lindsey, D. T. (2001). Direction repulsion in unfiltered and ring-filtered Julesz textures. Perception & Psychophysics, *63*, 226-240.
- Lorenceanu, J., Shiffrar, M., Wells, N., & Castet, E. (1993). Different motion sensitive units are involved in recovering the direction of moving lines. Vision Research, *33*, 1207-1217.
- Luck, S., & Vecera. (in press)., Stevens' Handbook of Experimental Psychology (Vol. 1). New York: Wiley.
- Macmillan, N. A. (in press). Signal Detection Theory, Stevens' Handbook of Experimental Psychology (Vol. 4). New York: Wiley.
- Magnussen, S., & Greenlee, M. W. (1992). Retention and disruption of motion information in visual short-term memory. Journal of Experimental Psychology: Learning, Memory & Cognition, *18*, 151-156.

- Marshak, W., & Sekuler, R. (1979). Mutual repulsion between moving visual targets. Science, *205*, 1399-1401.
- Mateeff, S., Genova, B., & Hohnsbein, J. (1999). The simple reaction time to changes in direction of visual motion. Experimental Brain Research, *124*, 391-394.
- Mather, G. (1980). The movement aftereffect and a distribution-shift model for coding direction of visual movement. Perception, *9*, 379-392.
- Mather, G., & Murdoch, L. (1994). Gender discrimination in biological motion displays based on dynamic cues. Proceedings of the Royal Society of London, Series B, *258*, 273-279.
- Mather, G., Verstraten, F. A. J., & Anstis, S. (1998). Preface. In G. Mather & F. Verstraten & S. Anstis (Eds.), The Motion Aftereffect: A Modern Perspective. (pp. vii-xii). Cambridge: MIT Press.
- McCarthy, J. E. (1993). Directional adaptation effect with contrast modulated stimuli. Vision Research, *33*, 2653-2662.
- McKee, S. P. (1981). A local mechanism for differential velocity detection. Vision Research, *21*, 491-500.
- McKee, S. P., Silverman, G. H., & Nakayama, K. (1986). Precise velocity discrimination despite random variations in temporal frequency and contrast. Vision Research, *26*, 609-619.
- McKee, S. P., & Watamaniuk, S. N. J. (1994). The psychophysics of motion perception, Visual detection of motion (pp. 85-114). New York City: Academic Press.
- Metzger, W. (1934). Beobachtungen ueber phaenomenale Identität [Observations on phenomenal identity]. Psychologische Forschung, *19*, 1-60.
- Morrone, M. C., Tosetti, M., Montanaro, D., Fiorentini, A., Cioni, G., & Burr, D. C. (2000). A cortical area that responds specifically to optic flow, revealed by fMRI. Nature Neuroscience, *3*, 1322-1328.
- Movshon, J. A., Adelson, E. H., Gizzi, M. S., & Newsome, W. T. (1986). The analysis of moving visual patterns. In C. Chagas & R. Gattas & C. G. Gross (Eds.), Pattern Recognition Mechanisms (pp. 177-151). New York: Springer-Verlag.
- Mulligan, J. B. (1992). Motion transparency is restricted to two planes. Investogative Ophthalmology and Visual Science, *33*, 1049.
- Nakayama, K. (1981). Differential motion hyperacuity under conditions of common image motion. Vision Research, *21*, 1475-1482.
- Nawrot, M., & Blake, R. (1991a). The interplay between stereopsis and structure from motion. Perception & Psychophysics, *49*, 230-244.
- Nawrot, M., & Blake, R. (1991b). A neural-network model of kinetic depth. Visual Neuroscience, *6*, 219-227.
- Nawrot, M., & Blake, R. (1993). On the perceptual identity of dynamic stereopsis and kinetic depth. Vision esearch, *33*, 1561-1571.
- Nawrot, M., & Sekuler, R. (1990). Assimilation and contrast in motion perception: explorations in cooperativity. Vision Research, *30*, 1439-1451.
- Neri, P., Morrone, M. C., & Burr, D. C. (1998). Seeing biological motion. Nature, *359*, 894-896.
- Newsome, W. T., Britten, K. H., & Movshon, J. A. (1989). Neuronal correlates of a perceptual decision. Nature, *341*, 52-54.
- Newsome, W. T., & Paré, E. B. (1988). A selective impairment of motion perception following lesions of the middle temporal visual area (MT). Journal of Neuroscience, *8*, 2201-2211.

- Nishida, S., Ashida, H., & Sato, T. (1994). Complete interocular transfer of motion aftereffect with flickering test. Vision Research, *34*, 2707-2716.
- Nishida, S., & Sato, T. (1995). Motion aftereffect with flickering test patterns reveals higher stages of motion processing. Vision Research, *35*, 477-490.
- Orban, G. A., de Wolf, J., & Maes, H. (1984). Factors influencing velocity coding in the human visual system. Vision Research, *24*, 33-39.
- Orger, M. B., Smear, M. C., Anstis, S. M., & Baier, H. (2000). Perception of Fourier and non-Fourier motion by larval zebrafish. Nature Neuroscience, *3*, 128-1133.
- Pack, C. C., & Born, R. T. (2001). Temporal dynamics of a neural solution to the aperture problem in visual area MT of macaque brain. Nature, *409*, 1040-1042.
- Pantle, A. (1973). Stroboscopic movement based upon global information in successively presented visual patterns. Journal of the Optical Society of America, *63*, 1280.
- Pantle, A. (1974). Motion aftereffect magnitude as a measure of spatio-temporal response properties of direction-selective analyzers. Vision Research, *14*, 1229-1236.
- Pantle, A. (1998). How do measures of the motion aftereffect measure up? In M. G & F. Verstarten & S. Anstis (Eds.), The Motion Aftereffect: A Modern Perspective (pp. 25-39). Cambridge: MIT Press.
- Pantle, A., Gallogly, D. P., & Piehler, O. C. (2000). Direction biasing by brief apparent motion stimuli. Vision Research, *40*, 1979-1991.
- Pantle, A., & Picciano, L. (1976). A multistable movement display: evidence for two separate motion systems in human vision. Science, *193*, 500-502.
- Parker, A. J., & Newsome, W. T. (1998). Sense and the single neuron: Probing the physiology of perception. Annual Review of Neuroscience, *21*, 227-277.
- Pascual-Leone, A., & Walsh, V. (2001). Fast backprojections from the motion to the primary visual area necessary for visual awareness. Science, *292*, 510-512.
- Pascual-Leone, A., Walsh, V., & Rothwell, J. (2000). Transcranial magnetic stimulation in cognitive neuroscience--virtual lesion, chronometry, and functional connectivity. Current Opinion in Neurobiology, *10*, 232-237.
- Pasternak, T. (1987). Discrimination of differences in speed and flicker rate depends on directionally-selective mechanisms. Vision Research, *27*, 1881-1890.
- Patterson, R. (1999). Stereoscopic (cyclopean) motion sensing. Vision Research, *39*, 3329-3345.
- Pelli, D., & Farell, B. J. (1999). Why use noise? Journal of the Optical Society of America A, *16*, 647-653.
- Pylyshyn, Z. W., & Storm, R. W. (1988). Tracking multiple independent targets: Evidence for a parallel tracking mechanism. Spatial Vision, *3*, 179-197.
- Qian, N., & Andersen, R. A. (1994). Transparent motion perception as detection of unbalanced motion signals. II. Physiology. The Journal of Neuroscience, *14*, 7367-7380.
- Qian, N., Andersen, R. A., & Adelson, E. H. (1994). Transparent motion perception as detection of unbalanced motion signals. I. Psychophysics. The Journal of Neuroscience, *14*, 7357-7366.
- Ramachandran, V. S., & Anstis, S. M. (1985). Perceptual organization in multistable apparent motion. Perception, *14*, 135-143.
- Raymond, J. E. (1993). Complete interocular transfer of motion adaptation effects on motion coherence thresholds. Vision Research, *33*, 1865-1870.
- Raymond, J. E. (2000). Attentional modulation of visual motion perception. Trends in Cognitive Science, *4*, 42-50.

- Raymond, J. E., O'Donnell, H. L., & Tipper, S. P. (1998). Priming reveals attentional modulation of human motion sensitivity. Vision Research, *38*, 2863-2867.
- Regan, D. (1989). Orientation discrimination for objects defined by relative motion and objects defined by luminance contrast. Journal of Experimental Psychology: Human Perception and Performance, *15*, 1389-1400.
- Riggs, L. A., & Day, R. H. (1980). Visual aftereffects derived from inspection of orthogonally moving patterns. Science, *208*, 416-418.
- Rizzo, M., & Nawrot, M. (1998). Perception of movement and shape in Alzheimer's disease. Brain, *121*, 2259-2270.
- Robson, J. G. (1966). Spatial and temporal contrast-sensitivity functions of the visual system. Journal of the Optical Society of America, *56*, 1141-1142.
- Rose, A. (1948). The sensitivity performance of the human eye on an absolute scale. Journal of the Optical Society of America, *38*, 196-208.
- Rudolph, K., & Pasternak, T. (1999). Transient and permanent deficits in motion perception after lesions of cortical areas MT and MST in the macaque monkey. Cerebral Cortex, *9*, 90-100.
- Salzman, C. D., Murasugi, C. M., Britten, K. H., & Newsome, W. T. (1992). Microstimulation in visual area MT: Effects on direction discrimination performance. Journal of Neuroscience, *12*, 2331-2335.
- Savelsbergh, G. J. P., Whiting, H. T. A., & Bootsma, R. J. (1991). Grasping tau. Journal of Experimental Psychology: Human Perception and Performance, *17*, 315-322.
- Schall, J. D. (2000). From sensory evidence to a motor command. Current Biology, *10*, R404-406.
- Schenk, T., & Zihl, J. (1997). Visual motion perception after brain damage: I. Deficits in global motion perception. Neuropsychologia, *35*, 1289-1297.
- Schiff, W. (1965). Perception of impending collision. Psychological Monographs, *79*, 1-26.
- Schrater, P. R., Knill, D. C., & Simoncelli, E. P. (2001). Perceiving visual expansion without optic flow. Nature, *410*, 816-819.
- Sekuler, A. B., & Bennett, P. J. (1996). Spatial phase differences can drive apparent motion. Perception & Psychophysics, *58*, 174-190.
- Sekuler, A. B., Sekuler, E. B., & Sekuler, R. (1990). How the visual system detects changes in the direction of moving targets. Perception, *19*, 181-196.
- Sekuler, A. B., & Sekuler, R. (1993). Representational development of direction in motion perception: a fragile process. Perception, *22*, 899-915.
- Sekuler, R., & Blake, R. (2001). Perception. New York: McGraw-Hill.
- Sekuler, R., & Ganz, L. (1963). A new aftereffect of seen movement with a stabilized retinal image. Science, *139*, 419-420.
- Sekuler, R., Sekuler, A. B., & Lau, R. (1997). Sound changes perception of visual motion. Nature, *384*.
- Shadlen, M. N., & Newsome, W. T. (1996). Motion perception: Seeing and deciding. Proceeding of the National Academy of Science USA, *93*, 628-633.
- Shulman, G. L., Ollinger, J. M., Akbudak, E., Conturo, T. E., Snyder, A. Z., Petersen, S. E., & Corbetta, M. (1999). Areas involved in encoding and applying directional expectations to moving objects. Journal of Neuroscience, *19*, 9480-9496.
- Sigman, M., Cecchi, G. A., Gilbert, C. D., & Magnasco, M. O. (2001). On a common circle: natural scenes and Gestalt rules. Proceeding of the National Academy of Science USA, *98*, 1935-1940.

Smith, A. T., & Edgar, G. K. (1991). The separability of temporal frequency and velocity. Vision Research, *31*, 321-326.

Smith, M. R. H., Flach, J. M., Dittman, S. M., & Stanard, T. (2001). Monocular optical constraints on collision control. Journal of Experimental Psychology: Human Perception & Performance, *27*, 395-410.

Snowden, R. J. (1990). Suppressive interactions between moving patterns: Role of velocity. Perception & psychophysics, *47*, 74-78.

Snowden, R. J. (1992). Sensitivity to relative and absolute motion. Perception, *21*, 563-568.

Snowden, R. J., Treue, S., Erickson, R. G., & Andersen, R. A. (1991). The response of area MT and V1 neurons to transparent motion. The Journal of Neuroscience, *11*, 2768-2785.

Snowden, R. J., & Verstraten, F. A. J. (1999). Motion transparency: Making models of motion perception transparent. Trends in Cognitive Sciences, *3*, 369-377.

Sperling, G., & Lu, Z.-L. (1998). A systems analysis of visual motion perception. In T. Watanabe (Ed.), High-level motion processing: Computational, neurobiological and psychophysical perspectives (pp. 153-186). Cambridge: MIT Press.

Spigel, I. M. (1960). The effects of differential post-exposure illumination on the decay of the movement after-effect. Journal of Psychology, *50*, 209-210.

Spigel, I. M. (1962a). Contour absence as a critical factor in the inhibition of the decay of a movement aftereffect. Journal of Psychology, *54*, 221-228.

Spigel, I. M. (1962b). Relation of MAE duration to interpolated darkness intervals. Life Sciences, *1*, 239-242.

Spigel, I. M. (1964). The use of decay inhibition in an examination of central mediation in movement aftereffects. Journal of General Psychology, *70*, 241-247.

Steiner, V., R. B., & Rose, D. (1994). Interocular transfer of expansion, rotation, and translation motion aftereffects. Perception, *23*, 1197-1202.

Stevens, S. S. (1951). Mathematics, Measurement and Psychophysics. In S. S. Stevens (Ed.), Handbook of Experimental Psychology (pp. 1-49). New York: John Wiley & Sons.

Stone, L. S., & Thompson, P. (1992). Human speed perception is contrast dependent. Vision Research, *32*, 1535-1549.

Sunaert, S., Van Hecke, P., Marchal, G., & Orban, G. A. (1999). Motion-responsive regions of the human brain. Experimental Brain Research, *127*, 355-370.

Sutherland, N. S. (1961). Figural aftereffects and apparent size. Quarterly Journal of Experimental Psychology, *13*, 222-228.

Symons, L. A., Pearson, P. M., & Timney, B. (1996). The aftereffect to relative motion does not show interocular transfer. Perception, *25*, 651-660.

Tanaka, K., & Saito, H. (1989). Analysis of motion of the visual field by direction, expansion/contraction, and rotation cells clustered in the dorsal part of the medial superior temporal area of the macaque monkey. Journal of Neurophysiology, *62*, 626-641.

Tayama, T. (2000). The minimum temporal thresholds for motion detection of grating patterns. Perception, *29*, 761-769.

Tetewsky, S. J., & Duffy, C. J. (1999). Visual loss and getting lost in Alzheimer's disease. Neurology, *52*, 958-965.

Thompson, P. (1982). Perceived rate of movement depends on contrast. Vision Research, *22*, 377-380.

Thompson, P. (1998). Tuning of the motion aftereffect. In G. Mather, Verstraten, F., & Anstis, S. (Ed.), The motion aftereffect: a modern perspective (pp. 41-55). Cambridge, MA: MIT Press.

Thorson, J., Lange, G. D., & Biederman-Thorson, M. (1969). Objective measure of the dynamics of a visual movement illusion. Science, *164*, 1087-1088.

Tong, F. (2001). Brain at work: Play by play. Nature Neuroscience, *4*, 560-561.

Tresilian, J. R. (1999). Visually timed action: time-out for 'tau'? Trends in Cognitive Science, *3*, 301-310.

Treue, S., & Maunsell, J. H. R. (1999). Effects of attention on the processing of motion in macaque middle temporal and medial superior temporal visual cortical areas. Journal of Neuroscience, *19*, 7591-7602.

Treue, S., & Trujillo, J. C. (1999). Feature-based attention influences motion processing gain in macaque visual cortex. Nature, *399*, 575-579.

Tse, P., Cavanagh, P., & Nakayama, K. (1998). The role of parsing in high-level motion processing. In T. Watanabe (Ed.), High-level motion processing: Computational, neurobiological and psychophysical perspectives (pp. 249-266). Cambridge: MIT Press.

Turano, K., & Pantle, A. (1989). On the mechanism that encodes the movement of contrast variations: Velocity discrimination. Vision Research, *29*, 207-221.

Tynan, P. D., & Sekuler, R. (1982). Motion processing in peripheral vision: Reaction time and perceived velocity. Vision Research, *22*, 61-68.

Ullman, S. (1984). Maximizing rigidity: The incremental recovery of a 3-D structure from rigid and non-rigid motion. Perception, *13*, 255-274.

Ungerleider, L. G., & Haxby, J. V. (1994). 'What' and 'where' in the human brain. Current Opinion in Neurobiology, *4*, 157-165.

Vaina, L. M., Cowey, A., & Kennedy, D. (1999). Perception of first- and second-order motion: separable neurological mechanisms? Human Brain Mapping, *7*, 67-77.

Vaina, L. M., Grzywacz, N. M., LeMay, M., Bienfang, D., & Wolpaw, E. (1998). Perception of motion discontinuities in patients with selective motion deficits. In T. Watanabe (Ed.), High-level motion processing: Computational, neurobiological and psychophysical perspectives (pp. 213-247). Cambridge: MIT Press.

Vaina, L. M., Lemay, M., Bienfang, D. C., Choi, A. Y., & Nakayama, K. (1990). Intact "biological motion" and "structure from motion" perception in a patient with impaired motion mechanisms: A case study. Visual Neuroscience, *5*, 353-369.

van den Berg, A. V. (1992). Robustness of perception of heading from optic flow. Vision Research, *32*, 1285-1296.

van Meeteren, A., & Barlow, H. B. (1981). The statistical efficiency for detecting sinusoidal modulation of average dot density in random figures. Vision Research, *21*, 765-777.

Van Oostende, S., Sunaert, S., Van Hecke, P., Marchal, G., & Orban, G. (1997). The kinetic occipital (KO) region in man: an fMRI study. Cerebral Cortex, *7*, 690-701.

Verghese, P., Watamaniuk, S. N. J., McKee, S. P., & Grzywacz, N. M. (1999). Local motion detectors cannot account for the detectability of an extended trajectory in noise. Vision Research, *39*, 19-30.

Verstraten, F., Fredericksen, R. E., & van de Grind, W. A. (1994). The movement aftereffect of bi-vectorial transparent motion. Vision Research, *34*, 349-358.

Verstraten, F. A., Fredericksen, R. E., van Wezel, R. J. A., Lankheet, M. J. M., & van de Grind, W. A. (1996). Recovery from adaptation for dynamic and static motion aftereffects: Evidence for two mechanisms. Vision Research, *36*, 421-424.

Verstraten, F. A. J., Fredericksen, R. E., van Wezel, R. J. A., Boulton, J. C., & van de Grind, W. A. (1996). Directional motion sensitivity under transparent motion conditions. Vision Research, 36, 2333-2336.

Viviani, P., & Stucchi, N. (1989). The effect of movement velocity on form perception: Geometric illusions in dynamic displays. Perception & Psychophysics, 46, 266-274.

von Grünau, M. W. (1986). A motion aftereffect for long-range stroboscopic apparent motion. Perception and Psychophysics, 40, 31-38.

Wade, N. J., Swanston, M. T., & De Weert, C. M. M. (1993). On interocular transfer of motion aftereffects. Perception, 22, 1365-1380.

Wade, N. J., & Verstraten, F. A. J. (1998). Introduction and historical overview. In G. Mather & F. Verstraten & S. Anstis (Eds.), The Motion Aftereffect: A Modern Perspective. (pp. 1-23). Cambridge: MIT Press.

Wallach, H., & O'Connell, D. N. (1953). The kinetic depth effect. Journal of experimental Psychology, 45, 205-217.

Walls, G. (1942). The vertebrate retina and its adaptive radiation. Bloomfield Hills, MI: Cranbrook Press.

Walsh, V., Ellison, A., Battelli, L., & Cowey, A. (1998). Task-specific impairments and enhancements induced by magnetic stimulation of human visual area V5. Proceeding of the Royal Society of London, Series B, 265, 537-543.

Warren, W. H., Jr, Blackwell, A. W., Kurtz, K. J., Hatsopoulos, N. G., & Kalish, M. L. (1991). On the sufficiency of the velocity field for perception of heading. Biological Cybernetics, 65, 311-320.

Warren, W. H., Jr, Blackwell, A. W., & Morris, M. W. (1988). Age differences in perceiving the direction of self-motion from optical flow. Journal of Gerontology, 44, P147-P153.

Warren, W. H., Jr, Kay, B. A., Zosh, W. D., Duchon, A. P., & Sahuc, S. (2001). Optic flow is used to control human walking. Nature Neuroscience, 4, 213-216.

Warren, W. H., Jr, Mestre, D. R., Blackwell, A. W., & Morris, M. W. (1991). Perception of circular heading from optical flow. Journal of Experimental Psychology: Human Perception & Performance, 17, 28-43.

Warren, W. H., Jr, Young, D. S., & Lee, D. N. (1986). Visual control of step length during running over irregular terrain. Journal of Experimental Psychology: Human Performance and Perception, 12, 259-266.

Watamaniuk, S. N. J. (1993). An ideal observer for discrimination of the global direction of dynamic random dot stimuli. Journal of the Optical Society of America A, 10, 16-28.

Watamaniuk, S. N. J., & Duchon, A. (1992). The human visual system averages speed information. Vision Research, 32, 931-941.

Watamaniuk, S. N. J., & McKee, S. P. (1995). 'Seeing' motion behind occluders. Nature, 377, 729-730.

Watamaniuk, S. N. J., & McKee, S. P. (1998). Simultaneous encoding of direction at a local and global scale. Perception & Psychophysics, 60, 191-200.

Watamaniuk, S. N. J., McKee, S. P., & Grzywacz, N. M. (1995). Detecting a trajectory embedded in random-direction motion noise. Vision Research, 35, 65-77.

Watamaniuk, S. N. J., & Sekuler, R. (1992). Temporal and spatial integration in dynamic random dot stimuli. Vision Research, 32, 2341-2347.

- Watamaniuk, S. N. J., Sekuler, R., & Williams, D. W. (1989). Direction discrimination in complex dynamic displays: The integration of direction information. Vision Research, *29*, 47-59.
- Watanabe, K., & Shimojo, S. (2001). When sound affects vision: effects of auditory grouping on visual motion perception. Psychological Science, *12*, 109-116.
- Watson, A. B. (1979). Probability summation over time. Vision Research, *19*, 515-522.
- Watson, A. B., Ahumada, A. J., Jr, & Farrell, J. E. (1986). Window of visibility: A psychophysical theory of fidelity in time-sampled visual motion displays. Journal of the Optical Society of America A, *3*, 300-307.
- Watson, A. B., & Robson, J. G. (1981). Discrimination at threshold: Labeled detectors in human vision. Vision Research, *21*, 1115-1122.
- Watson, A. B., Thompson, P. G., Murphy, B. J., & Nachmias, J. (1980). Summation and discrimination of gratings moving in opposite directions. Vision Research, *20*, 341-347.
- Watson, A. B., & Turano, K. (1995). The optimal motion stimulus. Vision Research, *35*, 325-336.
- Westheimer, G. (1988). Vision: Space and movement. In R. C. Atkinson & R. J. Herrnstein & G. Lindzey & R. D. Luce (Eds.), Stevens' Handbook of Experimental Psychology (Second ed., Vol. 1, pp. 165-193). New York: Wiley Interscience.
- Westheimer, G., & Wehrhahn, C. (1994). Discrimination of direction of motion in human vision. Journal of Neurophysiology, *71*, 33-37.
- Williams, D., Phillips, G., & Sekuler, R. (1986). Hysteresis in the perception of motion direction as evidence for neural cooperativity. Nature, *324*, Nov 20-26.
- Williams, D. W., & Sekuler, R. (1984). Coherent global motion percepts from stochastic local motions. Vision Research, *24*, 55-62.
- Wohlgemuth, A. (1911). On the after-effect of seen movement. British Journal of Psychology (Supp.), *1*, 1-117.
- Wuerger, S., Shapley, R., & Rubin, N. (1996). "On the visually perceived direction of motion" by Hans Wallach: 60 years later. Perception, *25*, 1317-1367.
- Wylie, D. R., Bischof, W. F., & Frost, B. J. (1998). Common reference frame for neural coding of translational and rotational optic flow. Nature, *392*, 278-282.
- Yantis, S. (1992). Multielement visual tracking: Attention and perceptual organization. Cognitive Psychology, *24*, 295-340.
- Yu, K. (2000). Can semantic knowledge influence motion correspondence? Perception, *29*, 693-707.
- Zacks, J. M., Braver, T. S., Sheridan, M. A., Donaldson, D. I., Snyder, A. Z., Ollinger, J. M., Buckner, R. L., & Raichle, M. E. (2001). Human brain activity time-locked to perceptual event boundaries. Nature Neuroscience, *4*, 651-655.