BIOLOGICAL IMAGE MOTION PROCESSING: A REVIEW

KEN NAKAYAMA

Smith-Kettlewell Institute of Visual Sciences, Medical Research Institute of San Francisco,
2232 Webster Street, San Francisco, CA 94115, U.S.A.

MOTION AS A FUNDAMENTAL VISUAL DIMENSION

Physics provides no special status for visual motion, skirting the issue as to whether it is fundamental or whether it is just the displacement of a visual image over time. Introspection is no more decisive. Is motion a basic phenomenological dimension like color and stereopsis, or is it derived, based on more primitive sensory processes, like space and time?

Color is an immediate experience. Likewise for stereopsis. Few fail to be impressed by the synthesis of solidity from two flat images in a stereoscope; the sense of depth is phenomenologically irreducible. With visual motion, however, there has always been the nagging doubt that it might not qualify as a fundamental sense, that it is reconstructed very late in our visual system or that it represents an elementary cognitive process. This view still persists and can be seen in the thesis that motion thresholds can be understood in terms of the memory of a position over time (Dimmick and Karl, 1930; Kinchla and Allan, 1969).

It is likely that the appreciation of motion as a fundamental biological sense was retarded by these alternative interpretations. Mounting evidence, accumulated over the past century and especially of late, however, leaves no doubt that motion is indeed a fundamental visual dimension. We start by providing a brief historical sketch of some of these independent sources of evidence, introducing some of the vocabulary of this review as well as highlighting some of the classical issues.

Perhaps the oldest demonstration of a separate motion process is the "waterfall" illusion or the motion aftereffect (MAE). Stationary objects are seen as moving in a direction opposite to that of previously viewed objects and the illusion is dramatic because a dissociation of motion and position is compelling. During the aftereffect, the target can be seen as moving, and yet it is also seen in the same position. This is a paradox unless one regards motion and position as separable sensory dimensions (Gregory, 1966). The illusion was noted by 19th century observers (Purkinje, 1825; Adams, 1834) and was first studied in detail by Wohlgemuth (1911). A modern variant of the motion aftereffect, and one that has led to some important quantitative observations, is that of direction specific (DS) adaptation (Sekuler and Ganz, 1963; Pantle, 1978). In this paradigm, luminance or contrast thresholds are selectively elevated after prolonged viewing of adapting targets moving in the same direction as compared to identical targets moving in the opposite direction.

Perhaps one of the most often quoted observations from early 20th century Gestalt psychology was from the study of classical apparent motion or "phi" by Wertheimer (1912). Movement or "phi" could be seen in response to two stationary flashes if these flashes were appropriately separated in space and time. The interpretation of "phi" was of particular significance because it suggested that a set of successively presented stationary images was equivalent to a "pure" motion stimulus, one that mimicked the encoding of real motion by the brain. Now we think that this view was an oversimplification and that the "phi" uncovered by Wertheimer is a very specialized process, now termed the "long range process". This is well suited to making comparisons over much greater distances and over longer time intervals than the mechanisms ordinarily used for the pick-up of real continuous motion (Anstis, 1980). Phi was an important discovery from a historical point of view, however, because it implied that motion was a primary sensory dimension, not capable of further reduction.

An important advance was made by German behavioral physiologists in the late 1950s and early 1960s. With an application of systems analysis to insect behavior, Reichardt (1961) and colleagues ascertained that the motion processing underlying the optomotor response in insects was mediated by local interactions between adjacent ommatidia or next-to-adjacent ommatidia of the insect compound eye. They also presented a computational model employing the principle of autocorrelation to account for the results (Reichardt, 1961). This work became a landmark, partly because it provided an explicit model of a complex visual process and partly because it was consistent with quantitative behavioral data. The model still remains as one of the most important general theories of motion processing and it has also opened the door for many other alternative formulations.

Single cell recording provided an entirely new source of evidence for the existence of motion sensitive mechanisms. Early pioneers of single unit recording found visual neurons which had a sensitivity to moving images (Barlow, 1953; Lettvin et al., 1959; Hubel and Wiesel, 1962, 1965). More detailed analysis of motion mechanisms came from a series of papers by Barlow and associates (Barlow and Hill, 1963; Barlow and Levick, 1965). They noticed that the impulse rate of certain cells were preferentially modified by changes in stimulus direction and velocity independent of other aspects of the stimulus such as contrast, shape and size.

Random dot stimuli have clarified an important distinction between apparent and real movement. Braddick (1974) found that if one displayed two uniformly textured random dot patterns in succession such that the dots in a central region were displaced as a unit from one frame to the next while the surrounding dots changed in an uncorrelated manner, the center dots would emerge as a vivid figure. This occurred only if the center displacement was less than a fixed amount (15 arc min in the original experiments). Braddick noted that this limit was much shorter than the limit for "phi" as described by Wertheimer (1912) and Korte (1915) and that this constituted evidence for a distinct short range motion

process. In contrast to "phi", this short range process did not work for dichoptic stimulation, thus emphasizing the conclusion that it is a very early process, occurring prior to the convergence of information from both eyes. As such, it indicated that the short range process was probably related to the emerging data from neurophysiological experiments.

Nakayama and Tyler (1981) also used random dots to isolate differential motion sensitivity from the contaminating effects of fine positional sensitivity seen in vernier acuity and other hyperacuity tasks (Westheimer, 1979). With this technique they were able to measure a minimum amount of displacement that could be detected and in this regard the experiments complement Braddick's paradigm which measured the maximum displacement. Nakayama and Tyler's work also revealed spatial characteristics of motion sensitivity which were very distinct from those underlying vernier acuity (see later section on D_{min}).

Because of these and many related observations (Sekuler, 1975), we conclude that image motion processing is a fundamental property of biological visual systems and that it can be experimentally isolated from other systems using a variety of experimental techniques. Color vision and binocular disparity detection have received wide interest, yet it is also clear that color processing is not present in all species and that binocular vision is restricted in animals with laterally placed eyes. As such, numerous animals either lack color vision or significant binocular vision or both. No animals have been shown to lack mechanisms for motion processing.

FUNCTIONAL BENEFITS OF IMAGE MOTION PROCESSING

What purpose might image motion processing serve? Perhaps a discussion of just seven possible roles will demonstrate that no single answer will suffice and that image motion sensitivity must play a very general and fundamental role in visual function.

(1) Encoding of the third dimension

The screen image on the retina is inherently two-dimensional and yet it is the job of the visual system to provide the third dimension, depth. Automatically, we think of stereopsis when we think of depth, but a moment's notice reminds us of many other "cues". One-eyed individuals or animals without significant binocular vision can navigate effectively in a complex three-dimensional environment. Of the so-called

---

*By relative distance is meant the ratio of distances between any two environmental points. This ratio can be calculated from the optical velocity field. Additional information such as the observers translational velocity or the distance of just one environmental point is sufficient to enable the reconstruction of absolute rather than relative distances (Nakayama and Loomis, 1974; Nakayama, 1983).
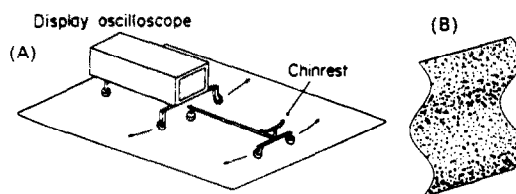

Fig. 1. Perceived depth obtained from differential image motion of random dots in a monocular image as reported by Rogers and Graham (1979). In (A), the observer places the head in the chin rest and is instructed to move the head from side to side. Differential shearing motion on the screen accompanies this head motion. (B) Provides a pictorial description of the depth percept. The observer sees a corrugated surface in depth.

monocular cues to depth, motion parallax is perhaps the most fundamental, because the optical velocity field contains the least assumption-laden information regarding the layout of the surrounding environment. Such a field contains rich information as to the slant of surfaces (Koenderink and Van Doorn, 1976; Longuet-Higgins and Prazdny, 1980) and the relative depth of surfaces from the observer (Nakayama and Loomis, 1974). In particular, if one were to translate in a rigid environment, it is theoretically possible to compute the relative* distances of environmental points with certainty.

Experimental confirmation of this theoretical viewpoint has been obtained by noting the depth seen in two-dimensional moving figures (Wallach and O'Connell, 1953; Rogers and Graham (1979; 1982). Figure 1(A) schematizes a situation where a monocular observer views a flat CRT containing a dense field of random picture elements (pixels). As would be expected the stationary dots on the screen appear flat with no depth. Rogers and Graham (1979) instructed observers to move their head from side to side in a chin rest. At the same time the observer's lateral motion was measured and was used to generate differential optical motion on the flat oscilloscope screen, mimicking the differential motion on the retina when viewing a three-dimensional corrugated surface. The result was a dramatic and unambiguous sense of depth, with the flat screen appearing as a corrugated sinusoidal surface in depth [see Fig. 1(B)]. The large local differences in perceived depth were as striking and compelling as that seen in random dot stereograms (Julesz, 1971). In addition, and important for our later discussion, the corrugated surface could be seen as absolutely stationary. Thus, motion or relative motion between parts of the random pattern were perceived as having pure depth with no motion. In general, the experimental findings of Rogers and Graham provide the motion counterpart to Wheatstone's (1838) synthesis of depth using binocular disparity or Julesz' (1960) random dot stereogram.

(2) Time to collision (TTC)

In the previous section we noted that the optical

velocity field provides information regarding the relative distances to environmental points. By itself. it does not supply absolute distance information. As such it would seem that the optical flow field would not contain sufficient information to supply the moving organism with an estimate of the "time to collision (TTC)" as it approached a visual target. Ordinarily, one would think that either observer velocity or target distance would be necessary.

A surprising mathematical result is that a time-to-collision parameter is available from the optical flow field even when absolute distances to the points are unknown (Hoyle, 1957; Lee, 1976). Furthermore, Schift and Detwiler (1979) and Todd (1981) have manipulated the rate of the dilation of images in animated figures and have found that time-to-collison can be perceived by human observers in the absence of information about distance and observer velocity.

Lower animals also seem to have the substrate to analyze these expanding optical flow fields. Flying insects, tethered and suspended from above, will extend their legs (the landing response) when exposed to a radially expanding pattern (Goodman, 1960; Braitenberg and Taddei-Ferretti, 1966). In addition there is a correlation between the velocity of the expanding array and the latency of this landing response (Coggshall, 1972).

(3) Image segmentation

Related to the problem of depth measurement is the need to parse the complex pattern of illumination in the optic array into different physical objects and to distinguish "figure" from "ground". Motion is eminently suited for this job because of the mathematical relation between neighboring points in the optical velocity field at the edges of objects. Points which are well within the boundaries of a visual object, for example, generally have the same or very similar velocities between neighboring points whereas this is not necessarily the case at the boundaries of objects in the image. This concept is mirrored by the well-known Gestalt principle of "common fate" where points moving at the same velocity are perceived as a coherent entity distinguishing it from background and other objects (Koffka, 1926).

In a more modern vein, Nakayama and Loomis (1974) have emphasized the properties of the velocity field at the boundaries of real objects, noting the existence of velocity discontinuities at these boundaries during translational motion of the head. To localize these velocity-defined "edges" they proposed the existence of concentrically organized receptive fields having a center–surround antagonism with respect to a particular velocity direction. Such units have been reported for many species (Sterling and Wickelgren, 1969; Collett, 1972; Bridgeman, 1972; Frost, 1978; Frost et al., 1981). In addition, Nakayama and Loomis (1974) hypothesized that the outputs of these specialized units, which were selec-

tive for particular directions of motion, were further summed across velocity direction by a "convexity" unit whose properties would provide a plausible mechanism to compute the location of edges of real objects independent of the direction of motion.

Recent neurophysiological studies on the characteristics of center/surround cells tuned to direction confirms the existence of this type of circuitry. Frost and Nakayama (1983) found that essentially all neurons recorded from nonsuperficial layers of pigeon tectum were sensitive to motion in the center and surround of their receptive fields such that the cell would only fire when motion in center and surround were in opposing directions. As suggested by Nakayama and Loomis (1974), this was a higher-order property, independent of movement direction. A similar relativity of motion sensitivity was also reported for some neurons in cat striate cortex (Hammond and Smith, 1982).

Allman et al. (1984) also found velocity cells with an antagonistic center–surround organization in the primate and its preferred velocity in the central portion of the RF was matched by the same preferred velocity tuning in the inhibitory surround. As such, these cells are insensitive to uniform motion over the center and surround and are highly sensitive to velocity differences between center and surround. A somewhat different motion mechanism was suggested to explain the figure–ground discrimination of the fly by Reichardt and Poggio (1979). They hypothesized that the outputs of neighboring movement detectors interact in a multiplicative-like fashion and then in turn, locally inhibit flicker detectors.

Whatever the underlying mechanism, the role of motion sensitivity for figural segmentation can be dramatic. It is seen in the motion parallax results of Rogers and Graham (1982) and it was the major phenomenon used by Braddick (1974) to uncover the short range motion process. It also enables insects to orient to a target of interest against an identically textured background (Virsik and Reichardt, 1976).

(4) Motion as a proprioceptive sense

Gibson (1954) suggested that visual motion was one of the primary sources of information for the moving organism to know about its own motion in relation to its environment (see also Turvey and Remez, 1979). Early work by Lee and associates measured the importance of optical flow information for postural control. Standing infants could be made to lose their balance and fall as a result of movement in the surrounding visual environment (Lee and Aronson, 1974). Environmental visual motion also destabilized the posture of adults, suggesting that visual motion information can override information obtained from stretch receptors in the limbs and gravity receptors in the inner ear (Lee and Lishman, 1975). Visual motion can also lead to a profound sensation of self-motion (vection), either as a rotation about a vertical axis or as a horizontal or vertical

translation (Brandt *et al.*. 1973; Johansson, 1971; Dichgans and Brandt. 1978).

The role of visual proprioception in lower animals has a long history. As an example, we note studies which have demonstrated compensatory torque and thrust responses from the wingbeats of flies. A major conclusion is that wingbeat amplitude can be appropriately modulated by different patterns of optical flow corresponding to rotation or translation of the animal with respect to different body axes (Gotz, 1968, 1972; Srinivasan, 1977). As an additional example, the optomotor response is thought to mediate the precise body alignment of fish during schooling (Shaw, 1978).

Single cell recordings in the rabbit support this proprioceptive role for vision, and provide an exciting hypothesis as to how such a system could be organized and complement other proprioceptive systems. Simpson, *et al.* (1981) made single unit recordings from the accessory optic system and the inferior olive of the rabbit and subjected the animal to whole field optical rotational along selected axes of rotation much like "stars" rotating around the eye in a "planetarium". By recording from many individual neurons, they established that there are essentially three cardinal axes to encode "whole field" optical rotation. A given single neuron will respond best when the whole velocity vector field is the optical equivalent of an observer rotating about one of three nearly orthogonal axes. One axis corresponds to rotation about a vertical axis and thus the corresponding visual neuron responds best to horizontal motion. The other two axes lie in the horizontal plane approximating 45° to the left and right of straight ahead (see Fig. 2).

The work of Simpson *et al.* (1981) suggests the existence of a close analogy between the three pairs of semi-circular canals, and these three classes of motion sensitive units. Each of the three pairs of semi-circular canals responds best to rotation along its rotational axis, an axis orthogonal to the plane of the canals. An analogous set of motion selective

Fig. 2. Representation of cardinal optical axes of rotation as suggested by single unit recordings in the rabbit (from Simpson *et al.*, 1981). These hypothetical axes are closely related to those measured for the semi-circular canals.

neurons appear similarly organized in terms of a system of discrete rotational basis vectors suitably arranged to encode the three-dimensional set of possible rotations of the organism. Thus motion information complements information obtained from the vestibular system and also appears to be organized in a similar fashion, even sharing the same set of coordinate axes.

At this point we speculate on the possible existence of yet another distinct set of whole field motion analyzers sensitive to 3 degrees of *translation*, thus providing the possible visual counterpart of the vestibular organs of saccule and utricle. A theoretical analysis of translational ego-motion, however, is burdened by the fact that distances of environmental objects affect the optical velocity field resulting from translation (Nakayama, 1983; Prazdny, 1981). This complication is not present for optical flow components associated with eye rotation.

(5) *Motion as a stimulus to drive eye movements*

Ever since the important experiments of Rashbass (1961), it has been recognized that the oculomotor pursuit system is driven by a velocity signal. Rashbass simultaneously stepped a visual target in one direction and initiated a constant velocity motion in the opposite direction. Thus position information (in the form of a step) was pitted against velocity information (in the form of a ramp). Surprisingly, the eye movement system responded separately to each, generating a smooth eye movement in response to velocity ramp and an oppositely directed saccade in response to the position step. Most often the smooth eye movement would precede the saccade, following the velocity of the motion even though this response increased the total positional error. Later work has suggested some additional contribution from a visual position encoding system (Robinson, 1965; Pola and Wyatt, 1980), but theoretical discussions of the oculomotor pursuit system still hinge directly on the notion that the visual system can indeed read velocity (Robinson, 1968).

The exact nature of image velocity coding in the pursuit system has been obscured by the fact that, ordinarily, pursuit is a closed loop system. Thus the actual extent to which the pursuit system receives a signal proportional to target velocity is not easily revealed by simply comparing stimulus velocity and smooth pursuit velocity because pursuit is also driven by an efference copy signal. To provide a direct examination of the visual velocity signals that drive the smooth pursuit system requires an opening of this feedback loop, either by retinal image stabilization or by careful measurement of the earliest portion of the smooth pursuit response (see Lisberger *et al.*, 1981; Lisberger and Westbrook, 1984; Kowler and McKee, 1984).

(6) *Motion as required for pattern vision*

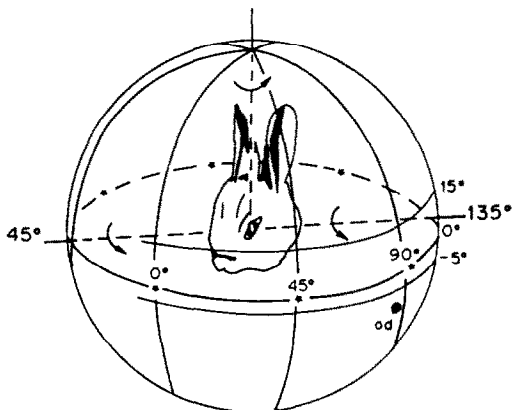Whether motion helps or hinders pattern vision has

been a puzzling and controversial question for decades, especially since the discovery that vision fades under retinal image stabilization (Riggs et al., 1953). Because of these results, it was thought that image motion must assist pattern vision by preventing fading of stabilized images. Surprisingly, a careful psychophysical examination of the problem indicated that visual acuity and vernier acuity were not affected by image stabilization (Keesey, 1960). So it seemed that motion had no major role in enhancing visual acuity.

Those concerned with the functions of the oculomotor system took the opposite point of view, however. They assumed that uncompensated retinal image motion must degrade vision, thus justifying the need for an oculomotor smooth pursuit and optokinetic system (Robinson, 1968).

From the recent work of Kelly (1979), it has become clear that both views are partly true. Kelly measured the contrast sensitivity of drifting and stationary sine wave gratings under retinal image stabilization. Motion indeed degrades the detection of spatial patterns having high spatial frequencies. For example, all information above 8 c/deg is lost if the pattern is moving at only $3°$/sec across the retina. Motion helps, however, in the pick-up of low spatial frequency information. Low spatial frequency sinusoidal gratings are essentially invisible when image motion is too slow or stabilized on the retina but they become visible when moving. Thus the naturally occurring movements of our retinal images do assist in the detection of low spatial frequency information, but are deleterious for the detection of high spatial frequency information. A description of this finding can be seen by looking ahead to Fig. 5(a).

(7) Image motion processing is useful for perceiving real moving objects

The sensing of the real motion of environmental objects is the most obvious use of image motion sensitivity and I have placed it last to emphasize its rather complex and shadowy relation to the perception of motion (see Gibson, 1968). The relation between retinal image motion and perceived motion is an interesting subject in its own right because there are clear examples where image motion leads to no sense of movement and, conversely, there are instances where no retinal image motion occurs and one sees lots of motion.

As mentioned earlier, the work of Rogers and Graham (1979) demonstrates very conclusively that even differential image motion need not lead to the perception of motion. Instead it can lead to a percept of pure depth with no sensing of absolute or differential motion. The converse can also be seen. A compelling sense of motion without retinal image motion is obvious when tracking one's visual afterimage in the dark (Yasui and Young, 1975), and the same occurs with stabilized visual images.

Of course, these examples do not mean that the

sense of motion is unrelated to retinal image motion. They do emphasize, however, that the conscious sense of motion is probably constructed rather late in the visual system and that it requires the combination of many different types of inputs. How much motion is seen as one moves one's head, for example, is proportional to the discrepancy between the perceived and the actual distance to a fixated target. This can be easily demonstrated in cases where a real three-dimensional figure (such as a 3-D Necker cube) is seen in an illusory reversed depth configuration (see Gregory, 1970), where the far side appears in front and the front side appears in back. During the reversed phase of the bi-stable depth percept, head motion leads to a dramatic reversal of the perceived motion, supporting a major role for perceived depth in the perception of motion. This view linking perceived motion to perceived depth has been outlined and supported in numerous papers by Gogel and colleagues (Gogel and Tietz, 1973; Gogel, 1980, 1982). An alternative and less convincing hypothesis is that perceived motion is proportional to the efference copy of associated pursuit movements just cancelling optokinetic nystagmus (Post and Leibowitz, 1982). This latter hypothesis can account for some limited aspects of absolute motion judgments but cannot explain the reversed motion phenomenon mentioned above, nor the illusory differential motion seen when an ordinary random dot stereogram is moved (Tyler, 1974).

In this review we make a determined effort to steer clear of the higher order issues of "perceived motion" and to emphasize early motion processing as a general purpose visual function having many beneficial roles in addition to the obvious task of perceiving the motion of real objects.

MULTIPLICITY OF FUNCTIONAL ROLES

From the foregoing partial list, it is clear that image motion processing has a large number of rather different roles to play in vision. The existence of these very diverse functions suggests that several motion systems might exist simultaneously. Each could have several functional roles and it is not impossible that one functional application could be served by more than one motion subsystem.

To provide rotational stabilization of the eye in space, for example, probably requires very different motion information than the task of image segmentation. In the latter case, a relatively high degree of retinotopic organization is required; whereas, in the former case, retinotopic mapping is superfluous but the system might be specialized to encode slow velocities over the whole visual field (Collewijn, 1982). This distinction is consistent with the existence of at least two types of motion systems in mammals, a cortical system devoted to analyzing motion at various loci in the visual field and a brainstem accessory optic system to analyze the average motion

over the whole field (Hoffman, 1982). Although separate, the two motion systems also show some evidence of interaction (Hoffman, 1982; Grasse et al., 1984), also showing a different ontogenetic timetable (Atkinson, 1979; Atkinson and Braddick, 1981; Naegele and Held, 1982).

The existence of these distinct functional roles for motion indicates that we must be careful when we lump together the results of many different types of experiments involving stimulus motion. With additional analysis it is expected that more motion systems will be delineated, showing profound differences both within and between species.

MOTION BLINDNESS?

Before describing the characteristics of motion sensitivity as revealed by some psychophysical and neurophysiological experiments, it is of interest to digress on a relevant clinical case study.

Selective sensory losses are well known. Many individuals lack either stereoacuity or color vision, yet otherwise have normal vision. It would seem that if motion were a specialized sense, a selective loss is a theoretical possibility. Such cases appear to occur very rarely for motion and this is fortunate for the victim because the results appear to be far more devastating than loss of color or stereopsis. Zihl et al. (1982) report a human case history which appears to be the clearest isolation of motion from other visual defects. CT scans indicated bilateral involvement of the parietal-occipital region. Vision was normal according to standard visual tests, including visual acuity, critical flicker frequency, color vision, and saccadic eye movement accuracy. For any task requiring the perception of movement, however, the subject was grossly deficient. Moving objects were seen as present at one locale and then another, but with little or no intervening movement.

It should be noted that the patient also lacked pursuit eye movements but had normal saccades, a result expected from the recognized importance of velocity signals for the pursuit system (Rashbass, 1961). In addition to a clear deficit in processing ordinary continuous motion, the subject was also insensitive to stimuli that ordinarily elicit classical apparent movement or the "long range" process. Isolated pairs of point stimuli placed at spatial distances and temporal intervals larger than that required for the short range process but ordinarily appropriate for the long range process were not seen as moving. This suggests that the brain damage occurred where the "long range" and the "short range" process to encode continuous movement are combined or that the damage affected each system separately.

The patient had a number of additional deficits which suggest the importance of visual motion sensitivity for a variety of brain functions. Foremost was a deficiency in self-locomotion and the processing of real moving objects. Crossing the street in the face of on-coming traffic was frightening because a car would seem far away at one instant then suddenly it was dangerously close. Pouring a full cup of coffee without having it overflow was especially difficult, perhaps reflecting the importance of visual velocity to extrapolate future events in time.

The social consequences of the brain damage were serious. Paricularly disconcerting was the situation where the patient would be conversing with another individual and where a third person would enter the room and go unnoticed. Then suddenly the patient would see the new person. This latter example indicates an important orienting and attentive role for motion in addition to the seven that were outlined above. Not having this capacity, the observer would be forced to rely on cognitive faculties, remembering that the room was empty and possibly scanning the entrances for such events. The patient also had difficulty in conversing with others because of an inability to read facial expressions. This suggests that the dynamic component of facial expression is more important than one might have thought and that motion processing could play a major role.

Because the neuronal damage could have affected other brain areas in addition to those responsible for motion, it is impossible to be sure that these deficits are due to a loss of motion itself and not the result of the loss of some other functional system. The pattern of dysfunction is consistent with such a loss, however.

Laboratory experiments subjecting normal humans to reduced cue environments where motion sensitivity is abolished provides some interesting corroboration of these clinical observations. For example, stroboscopic environments having interflash intervals greater than the interval required for the short range motion process suggest similar deficits. Size constancy is dramatically attenuated under strobe illumination, showing its greatest reduction for a strobe rate of about 8 Hz (Rogowitz, 1983). Furthermore, the sense of observer self-motion (vection) is also reduced for some rates of stroboscopic illumination (Schor and Narayen, 1981).

PARALLEL AND SERIAL PROCESSING WITHIN AN EARLY MOTION PROCESSING SYSTEM: A SKELETAL MODEL

Up until now we have raised the possibility of several motion systems and their various functional roles. Ignoring this complexity for the moment, we focus on the possible constituent components *within* a motion system, especially the "short range process" as revealed by psychophysical experiments in humans and from single unit recordings in the geniculo-striate system of primates.

At the outset it should be recognized that even this reduced problem is formidable. To get some guidance it may be worthwhile to reflect on progress made in one of the most conceptually developed areas of

physiological optics, namely color. After prolonged debate, the originally conflicting concepts of trichromacy and chromatic opponency have found resolution by postulating two very different but compatible stages. The three cone types define the basic tridimensionality of the system at the receptors and the immediate recoding of their outputs forms a second stage of opponent colors. Thus chromatic processing can be seen in terms of a system of parallel channels undergoing serial transformation.

Although motion is bound to be different from chromatic processing in its essentials, its circuitry is also likely to contain parallel and serial elements of comparable complexity. The neuroanatomy of the visual pathway with its converging and diverging connections certainly points in this direction.

Such complexity, however, makes it difficult to interpret psychophysical results. In studies of motion, no less than in color, psychophysical data can be influenced by many elements and stages. It can reflect the properties of individual parallel system elements within a given stage or between stages. Given the difficulty of interpreting the data as well as the growing number of recent studies on motion processing, an acute need for a common conceptual framework becomes apparent.

A large number of motion theories do exist and several distinct classes of algorithms to encode direction and velocity have been suggested (Reichardt, 1961; Foster, 1971; Marr and Ullman, 1981, Barlow and Levick, 1965; van Santen and Sperling, 1984). These and others will be described later in a section on Computational Theories. Despite fundamental differences there is surprisingly good agreement, at least implicitly, between many of these theories with respect to some overall organizational aspects of the

motion system. This common overview deserves attention because it can organize data from a wide variety of experimental paradigms.

A pictoral outline of this overview can be seen in Fig. 3. At the front end of this picture are a series of input receptive fields (RFs), each sensitive to both the spatial position of the retinal image and also to different spatial frequency ranges (Blakemore and Campbell, 1969; Campbell et al., 1969; Wilson and Bergen, 1979; Marcelja, 1980). As such, these units code both local sign (position) and spatial frequency (size or image scale). Particular theories provide rather explicit circuitry to do the motion sensing (see later). We simplify this by thinking of a pair of receptive fields spaced by an effective distance Δs and an equivalent time Δt which delays the signals before their combination. See Fig. 3(B). For the moment the two-dimensional shape of the spatial receptive fields is left unspecified.

Following the input RFs is a stage of directional sensitivity (DS) and velocity sensitivity which operates on the signals coming from the input RFs such that the outputs of the DS sub-units are highly dependent on the direction and/or the velocity of the stimulus. Such units are generalizations of the directional processing sub-units proposed by other models, i.e. the comparator following the Reichardt multipliers (Reichardt, 1961) or the directionally selective sub-unit proposed by Barlow and Levick (1965).

Following the DS sub-units is a hypothetical stage of spatial and temporal integration. The spatial integration concept originated in a model suggested by Barlow and Levick (1965) and the temporal integrator have been incorporated into many previous models (Reichardt, 1961; Foster, 1971). The stage of spatial and temporal integration is lumped for simplicity only. No evidence or theory dictates a separation as yet.

Given this skeletal model, we are in a better position to ask some systematic questions about the short range motion process at several different levels.

Nondetection of moving gradients supports a spatial frequency bandpass front end and precludes gray level encoding

Ultimately a motion detecting system must rely on some difference in image radiance to sense motion. It is obvious that a moving surface without any variation in radiance cannot be seen as moving. Are there additional surfaces that cannot be seen to move? One surface which has a wide range of luminance values is that of a linear gradient.

Before discussing the perception of moving linear gradients, we note a simple mathematical relation between illumination and velocity. Given constant illumination it can be shown that for any nonuniform luminance distribution:
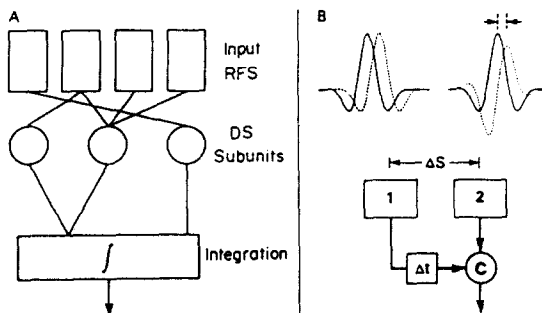
Fig. 3. (A) Overview of a proposed skeletal model of motion processing, showing several stages. The first stage consists of elements with receptive fields (RFs) which are influenced by the position and the spatial frequency of the stimulus. The second stage consists of directionally selective (DS) subunits which encode velocity and direction. A final stage consists of a spatial and temporal neural integrator. (B) shows more details of the hypothetical DS mechanism to detect rightward moving patterns. Pairs of receptive fields, either a displaced pair of symmetric detectors or a symmetric–antisymmetric pair combine at C through a circuit which effectively delays the left RF signal by Δt. The exact form of the combination at C (multiplicative or additive) is left unspecified (but see later section on Computational Theories of Motion Processing).

$$V_x = -(dI/dt)/(dI/dx) \qquad (1)$$

where V_x is the velocity in the x direction and dI/dt and dI/dx represent local temporal and spatial image intensity gradients (Hadani et al., 1980; Horn and Strunk, 1980; Fennema and Thompson, 1979). An explicit biological model based on this mathematical relationship would predict that the motion of any nonuniform luminance distribution could be seen. Such a model implies that one could compute velocity by taking a temporal derivative of the luminance at a point and dividing it by the spatial derivative at the same point.

Although the scheme is mathematically correct, it does not appear to be applicable to biological image processing because of the results obtained for the special case of a linear luminance gradient. First, spatial gradients are poorly sensed by the visual system (McCann et al., 1974). Second, Nakayama and Silverman (1983) presented a steep linear spatial luminance gradient on a CRT face and moved the light distribution in a direction along the gradient. Although equation (1) allows for the recovery of movement signals, no movement was seen even for displacements more than an order of magnitude greater than for ordinary motion thresholds. It appears that movement information is not available from direct operation on the absolute intensity level or by a combination of a first spatial and temporal derivative. That motion detection cannot be seen in this moving pattern is in accord with the skeletal model proposed in Fig. 3. Bandpass spatial input filters are rather insensitive to linear gradients of luminance or absolute levels of luminance.

There is also an interesting ecological reason why the motion system might not operate on gray levels. To do so would invite contamination by extraneous changes in ambient illumination levels. To appreciate this point, think of the consequences of having the sun go behind a cloud. For some surfaces this ambient change would yield a similar change as that associated for movement of the surface itself (see Fig. 4).

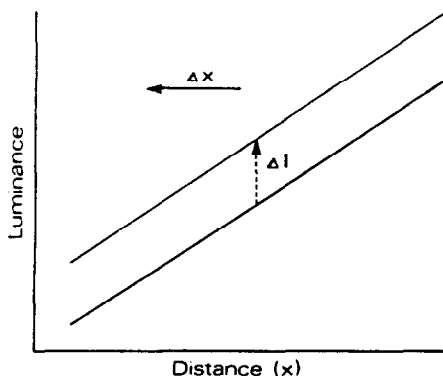The foregoing considerations suggest that the earliest stages of motion encoding must operate on

Fig. 4. Luminance profile of a gradient stimulus which is moved to the left by amount denoted by the vector ΔX. An equivalent representation of this movement is an increase of luminance as represented by the arrow labeled ΔI.

something other than the raw luminance distribution. In fact, nearly all recent models of motion encoding suggest sets of differently tuned spatial frequency bandpass filters preceding the stage of velocity encoding (van Santen and Sperling, 1984; Marr and Ullman, 1981; Watson and Ahumada, 1983). Of interest is to consider both the spatial and temporal characteristics of these input filters to determine the relationship of these input filters to those involved in the earliest encoding of contrast. Numerous paradigms have been used to determine these and other characteristics of the motion system and it is no trivial matter to compare the results of the various paradigms. Because random dot stimuli have the ability to isolate motion sensitivity rather directly, we treat experiments using these stimuli first. Then we will proceed to discuss results obtained using sinusoidal gratings because these stimuli have particular advantages in analyzing the spatio-temporal characteristics of motion processing.

RANDOM DOT STIMULI

D_{max}

Random dots were employed in the study of motion sensitivity shortly after their introduction by Julesz. Among the earliest studies were experiments by Bell and Lappin (1972) and Pollack (1972).

The use of random dots attracted widespread attention only after Braddick's (1974) original experiments, however, where he measured the maximum displacement (D_{max}) that could be seen as moving coherently. His stipulation of a short range process with a maximum finite distance of approximately 15 arc min was consistent with the concept of a finite size receptive field sub-unit already postulated by computational (Reichardt, 1961) as well as neurophysiological models (Barlow and Levick, 1965). The work was also important insofar as it distinguished the concept of a short range process from "classical" apparent movement, effectively focusing the discussion on more physiological hypotheses (Anstis, 1978; Braddick, 1980). The distinction had been suggested earlier (Gregory, 1966) but two experimental observations were significant. The short range process did not work under dichoptic presentation (presenting one pattern to one eye and the next to the other) and it could be seen only for rather short inter-stimulus asynchronies, much shorter than had been seen for classical apparent movement.

Braddick indicated that the short range process had a fixed D_{max} independent of dot density. This view, however, was disputed by Lappin and Bell (1976) who determined that the dot spacing was of importance in determining the maximum distance that could be spanned. Later experiments by Baker and Braddick (1982), varying field size of the moving random dot pattern indicated that this maximum distance (D_{max}) could deviate significantly from 15 arc min, being much smaller for small field sizes and

larger for much larger field sizes. They suggested that a determining factor was retinal locus, with eccentric postitions being capable of encoding a larger D_{max}. These general observations were also confirmed by Petersik et al. (1983) and Nakayama and Silverman (1984). Furthermore it was shown that the spatial frequency content of the random dot pattern was important in determining D_{max} (Chang and Julesz, 1983; Nakayama and Silverman, 1984). Bandpass patterns with high spatial frequency content had much lower values of D_{max} than patterns with lower spatial frequency content. Chang and Julesz also found that spatial filtering of random dots patterns in a direction orthogonal to the direction of motion had the least effect in reducing D_{max}.

D_{min}

In addition to D_{max}, it is also of interest to measure the minimum amount of motion (D_{min}) that can be detected. This measurement has been made on numerous occasions but the interpretation has been obscured by the problem of disentangling motion sensitivity from an awareness of changing position over time. Graham et al. (1948), for example, measured the ability of observers to detect differential motion in a vernier line stimulus with the upper and lower section moving in opposite directions. Although reliable thresholds were obtained, it was unclear to the authors whether they were measuring sensitivity to motion or a sensitivity to a change in position. Human observers are capable of noting the presence or absence of a vernier offset at the beginning or the end of a moving stimulus and inferring motion. An analogous question occurs when observing a clock. Are we really using our motion detecting system to appreciate the motion of the minute hand or are we constructing this from our memory of position over time?

Nakayama and Tyler (1981) presented a random dot pattern to an observer and moved it differentially in a horizontal direction such that its motion could be defined as a standing wave of shearing motion with a defined spatial and temporal frequency. By spatial frequency, we mean movement spatial frequency such that the instantaneous velocity of any given row was a spatial sinusoidal function of its vertical position. It was expected that there would be no codable position cues in this highly dense random pattern (see Attneave, 1954) and the observer's thresholds should be reflections of motion rather than position sensitivity. Nakayama and Tyler (1981) confirmed this hypothesis by measuring D_{min} (the minimum motion threshold) as a function of temporal frequency using a variety of spatial

configurations. including random dots and lines. When easily codable position cues were absent as in the case of random dots or in spatial configurations where static hyperacuity is poor (Tyler. 1973; Westheimer and McKee, 1977), the threshold was determined by the maximum velocity in a sinusoidally varying displacement profile. When position cues were present. however, the thresholds were determined by the size of the displacement alone, independent of velocity.

Nakayama and Tyler (1981) further isolated motion from position sensitivity by showing that they possess very different spatial characteristics. They measured the spatial frequency* dependence of shearing motion thresholds in random dots and compared it with that of periodic vernier acuity (Tyler, 1973). Although the best thresholds were about the same, as low as 5 arc sec in each case, the dependence on movement spatial frequency was entirely different. Static hyperacuity was best at about 2–3 c/deg whereas shearing motion sensitivity was best below about 0.6 c/deg and remained at the optimum sensitivity for spatial frequencies as low as 0.15 c/deg. Thus motion sensitivity falls off above a spatial frequency of 0.6 c/deg which corresponds to a half period of approximately 50 arc min. Similar findings have also been reported in a study comparing shearing sensitivity in monkey and human, also showing a low frequency rise in threshold for very low spatial frequencies of motion (Golumb et al., 1985). An increased area of spatial integration for motion can also be seen in the data of studies using moving line or moving single dot stimuli (Westheimer, 1979; Legge and Campbell, 1981).

One additional property distinguishes the sensitivity to small amounts of relative motion from fine positional sensitivity. Differential position sensitivity is surprisingly immune to overall image motion such that vernier thresholds are unaffected for velocities up to $3°/sec$ (Westheimer and McKee, 1975). Differential motion sensitivity, on the other hand, is severely disrupted by common image motion. Velocities of $3°/sec$ can raise differential motion thresholds by an order of magnitude (Nakayama, 1981). All of these distinguishing features of motion sensitivity using random dots indicate that motion cannot be derived from psychophysically measured characteristics of position sensitivity. This underscores the view that motion is a unique visual subsystem.

Intermediate values of velocity

So far we have discussed only the upper and lower spatial limit of motion encoding. To examine the characteristics of motion between the lower and the upper threshold, an alternative strategy is to choose an intermediate velocity and to manipulate the visibility of coherently moving random dots, usually by varying their luminance or contrast (Ball and Sekuler, 1980).

In this regard. a systematic approach was taken by

---

*Spatial frequency used here refers to movement spatial frequency, the number of cycles per degree of shearing movement. It does not refer to the spatial frequency of the luminance modulation, the more common usage.

Van Doorn and Koenderink (1982a,b) in their spatio-temporal characterization of the motion system at different velocities. They fixed the total r.m.s. contrast of random dots to a high suprathreshold level and electronically varied the ratio of contrast of coherently moving random dots to the contrast of incoherently moving dots, defining signal-to-noise ratio (SNR) as the square of this contrast ratio. The observers task was to increase SNR until coherent motion was just perceived. Then they made two novel manipulations, one temporal and one spatial.

In the temporal case the coherently moving dots moved as a unit over the whole field and were reversed in the direction of motion every Δt msec. Depending on the Δt and the velocity, three qualitatively different percepts were apparent. At very long intervals (300 msec or more), the target was seen as reversing direction over time and the threshold SNR was no different than for the condition of no reversal. At very short intervals, say 10 or 20 msec, no reversals of motion were seen. Instead, the observer saw transparency, with two planes of dots moving in opposite directions. For some restricted intermediate values of Δt, the percept of coherent motion disappeared or became extremely weak with threshold SNR rising dramatically. This critical value of Δt where motion became much weaker decreased systematically with increasing velocity [look ahead to Fig. 6(B)].

In the spatial case, Van Doorn and Koenderink (1982b), set up a number of equally spaced horizontal rows where the vertical motion in each alternate row was moving in the opposite vertical direction. In a situation rather analogous to the temporal case, they also saw three distinct percepts, depending on Δs (the width of these panels). For very large panels, the observer simply saw alternative panels of random dots moving in opposite directions. For very small panels (on the order of 2 arc min), the observer again saw two transparent planes moving in opposite directions. For intermediate panel widths, however, the percept of coherent motion either disappeared or became very weak accompanied with a sharp rise in threshold SNR. This critical value of Δs increased systematically with velocity [see Fig. 6(A)].

Van Doorn and Koenderink (1982 a,b) argue that these critical spatio-temporal intervals provide an estimate of the fundamental spatio-temporal parameters of the input stages of the motion system as pictured by models schematized in the inset of Fig. 3. The fact that they varied systematically for different velocities indicates that many different motion sensors exist in any given retinal region and that these different units must examine the characteristics of the image simultaneously.

Nakayama and Silverman (1984) were also able to relate spatial and temporal properties of motion in relation to velocity using random dots in a two step displacement paradigm. They estimated an upper displacement limit (D_{max}) associated with a maximum

perceptible velocity (V_{max}). They found that D_{max} increased with increasing velocity [see Fig. 6(A)].

EXPERIMENTS USING SINUSOIDAL GRATINGS

Up until now we have restricted most of our discussion to experiments using random dots. Although they are well suited to isolate motion processing from other forms of visual sensitivity, they have disadvantages. As they contain a wide range of spatial frequencies they do not isolate mechanisms tuned to particular spatial frequencies or scales in the image. As such they may have some limitations in delineating the nature of the fundamental circuitry. A complementary approach is to use sine wave gratings, stimuli having the greatest ability to isolate mechanisms having the same receptive field sizes.

Before proceeding, it is important to ask whether the short range motion system as isolated by sine wave stimuli is the same motion system stimulated by random dots. An experiment by Green and Blake (1981) provides support for this view. They found that by displacing a low spatial frequency grating (0.5 c/deg) by a phase shift of $90°$, a clear percept of motion was seen under normal viewing. In dichoptic viewing, however, motion could not be identified. This lack of motion under dichoptic presentation mirrors the original results obtained by Braddick (1974) using random dots and supports the view that sine wave gratings stimulate the same motion system as probed by random dots.

One of the earliest approaches using sinusoidal gratings to examine "short range" motion sensitivity was the measurement of motion aftereffect (MAE) strength as a function of the temporal and spatial frequencies in the adapting gratings. Pantle (1974) found that the velocity of the adapting stimulus provided an inadequate account of the data. If one looked over the whole range of spatial frequencies used to elicit the MAE, there was no single velocity which was optimal. Instead it appeared that the best velocity to obtain adaptation shifted downward with increasing spatial frequency such that it was more accurate to say that the system became most adapted for a preferred value of temporal frequency for almost any spatial frequency. Using low photopic luminance levels, Pantle found that a temporal frequency of 5 Hz elicited the largest MAE. A similar result using directionally selective adaptation was reported by Tolhurst (1973). This was puzzling. The directionally selective adaptation paradigm revealed directionally selective mechanisms but the adaptation was determined mainly by the temporal frequency of the moving grating, not velocity.

Measurement of the contrast sensitivity of moving gratings is a very different paradigm than that used to measure the MAE or directionally selective adaptation. Suprisingly, such experiments provided a similar set of results. This was evident in an early comprehensive study of drifting sine wave gratings at
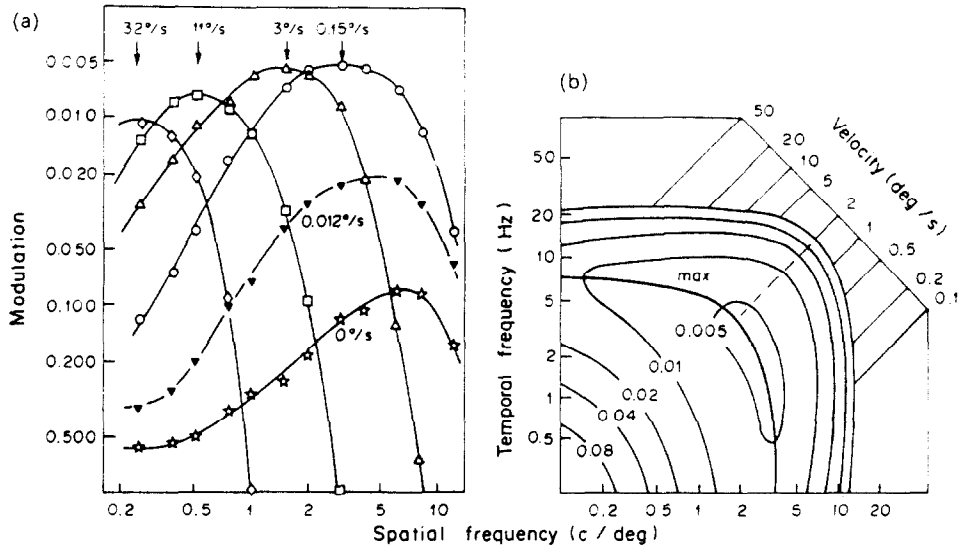
Fig. 5. (a) Contrast sensitivity functions for the detection of drifting sine wave gratings under conditions of stabilized viewing. Each curve represents a different drift velocity as labeled (from Kelly, 1979). (b) Contour map of an estimated spatio-temporal threshold surface derived from data shown in (b). Each contour line represents equal increments in log threshold as labeled. The velocity axis is along the $-45°$ line such that equi-velocity lines are parallel $+45°$ lines in this log-log representation (from Kelly, 1979).

different velocities (Watanabe *et al.*, 1968). More recently Kelly (1979) measured the same spatial contrast sensitivity functions for drifting gratings under conditions of retinal image stabilization. For a given velocity the contrast sensitivity function had a very similar shape to any other velocity, and with the exception of very low velocities (below 0.15 deg/sec), the contrast sensitivity function was shifted primarily along the spatial frequency axis depending on velocity [Fig. 5(a)]. Spatial frequency tuning was shifted towards lower spatial frequencies for increasing velocities. Plotting the same data in terms of temporal frequency yielded nearly overlapping curves peaking around 8 Hz. By making a few curve fitting assumptions from the data of the form seen in Fig. 5(a), Kelly constructed a two-dimensional threshold surface indicating sensitivity to drifting gratings over a wide range of spatial and temporal frequencies [see Fig. 5(b)]. Because of the mathematical relationship of velocity to spatial and temporal frequency [see equation (2) in the section on Fourier Description on Moving Images], loci of constant velocity in the spatial frequency (SF) and temporal frequency space (TF) are plotted as a set of parallel lines oriented at 45° in a log TF, log SF representation as shown in Fig. 5(b).

It should be noted that in Kelly's experiment observers were asked to detect the presence of the pattern, not to detect motion or direction. So on the

face of it, the experiment is not about motion sensitivity but about pattern sensitivity. The fact that different spatial frequency components are most visible at different velocities, however, suggests a close relation between pattern detection and velocity.

Burr and Ross (1982) modified Kelly's experiment to examine directional selectivity. Like Kelly they also varied the spatial frequency of drifting gratings moving at various velocities, plotting spatial contrast sensitivity functions for selected velocities, including some very high velocities. The study differed from Kelly's because the observer was asked to see the direction of motion and not the mere presence of the grating. Despite the difference in task, the results were in essential agreement with that of Kelly (1979)*. The curves slide horizontally to lower spatial frequencies with increasing velocity. Furthermore the curves for higher velocities (greater than 10°/sec) nearly coincided with a peak near 10 Hz when plotted as a function of temporal frequency.

These seemingly very different visual phenomena: MAE, pattern detection and direction discrimination, fit a similar set of functions. Maximum fatigueability, sensitivity and discriminability occur when motion stimuli are in the range of 5–10 Hz, with a tendency towards a lowering of the optimal temporal frequency for the slowest velocities and a decrease in the optimal spatial frequency at high velocities.

---

*Congruence between detection and discrimination of direction thresholds is evident for a wide range of velocities. For very low velocities, however, there appears to be a dissociation between these two thresholds. Discrimination requires more contrast as velocity is diminished (Watson *et al.* 1980; Mansfield and Nachmias, 1981; Green, 1983).

A COMMON SPACE-TIME FRAMEWORK TO ACCOUNT FOR SINUSOIDAL AND RANDOM DOT DATA

One of the challenges in the psychophysical examination of motion sensitivity is to provide an overall framework to account for the diverse results obtained

under a wide variety of experimental paradigms. Of interest is to reconcile work done with random dots with that obtained using sinusoidal grating stimuli. Each type of experiment reveals a spatial and temporal dependency with respect to velocity and as a consequence we propose a theoretical overview to account for both. It is suggested that by making a few reasonable assumptions, one can use spatio-temporal frequency data to make an estimate of the spatio-temporal characteristics of the early motion processing system, i.e. a description of the signal processing which occurs in the vicinity of the RFs and the DS subunits as indicated in Fig. 3. Then having this estimate, we can compare it to estimates made using nonperiodic stimuli, i.e. random dots.

To make this estimate from sine wave data, we assume that the motion system is fed by spatial frequency tuned receptive fields spaced so that they are in quadrature phase [see right hand pair of receptive field profiles in Fig. 3(B)]. This orthogonal basis set of receptive fields has been suggested for the visual cortex by Marcelja (1980) and corresponding single unit evidence for the existence of orthogonal sine and cosine cells has been reported by Pollen and Ronner (1979). This is a reasonable assumption because an orthogonally configured array of such channels optimally codes visual information with the smallest number of channels. It provides an encoding scheme for the visual cortex which maximizes signal-to-noise ratios given a fixed number of neurons (Sakitt and Barlow, 1982). Such a configuration of input filters has also been suggested in a modified version of the Reichardt model (van Santen and Sperling, 1984). Psychophysical results consistent with this spatial phasing comes from studies by Nakayama and Silverman (1985). They measured contrast sensitivity for the detection of motion direction for gratings which were instantaneously displaced by various angles of spatial phase. For a wide range of spatial frequencies, the peak contrast sensitivity coincided with a phase shift of 90 degrees and the contrast sensitivity closely followed an expected sine function with respect to phase angle.

Assuming that motion is detected by pairs of spatial frequency filters in quadrature phase, we suggest that the optimal equivalent Δt between these two input detectors can be inferred. It corresponds to one quarter of the time period defined by the best temporal frequency because any given spatial frequency component will optimally stimulate this pair of detectors if this time interval is selected. These optimal time periods are plotted as function of velocity in Fig. 6(B). The solid curve is obtained by taking the peak sensitivity estimated from Kelly (1979) which we reproduced as Fig. 5(b) and replotting this peak in terms of time vs velocity coordinates. The locus of this line in Fig. 6(B) indicates that the optimal temporal interval starts high for slow velocities and then decreases to an asymptotic limit for higher velocities. This is also confirmed by calcu-
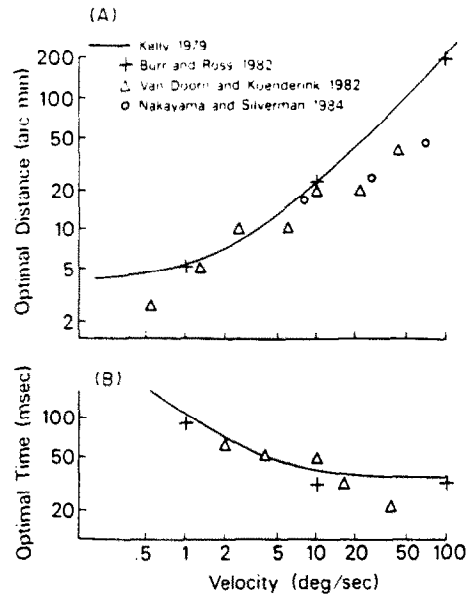

Fig. 6. Optimal temporal and spatial intervals for the inputs of hypothesized DS subunits [see Fig. 3(B)] as estimated from diverse experimental paradigms using drifting gratings and random dots. The solid line comes from noting the peak spatial and temporal frequency contrast sensitivity for detection as plotted in Fig. 5(b) (from the drifting sinewave data of Kelly, 1979) and calculating the optimum spatial and temporal intervals as described in the text. Crosses come from data measuring the contrast sensitivity for direction discrimination (Burr and Ross, 1981) analysed in the same way as the detection data of Kelly (1979). The solid circles are derived from entirely different experiments requiring the observer to see coherent motion of random dots in a field of dynamic visual noise (Van Doorn and Koenderink, 1982), and the solid squares are derived from data where the Braddick upper limit is measured as function of temporal parameters (Nakayama and Silverman, 1984). It should be noted that despite wide differences in the experimental paradigm and observer task, estimates of optimal spacing and timing for a given velocity show considerable similarity.

lating the peak timing (equivalent to one quarter of the reciprocal of the best temporal frequency) obtained by Burr and Ross (1982) and these are plotted as the crosses in Fig. 6(B). It should be noted that the results of Burr and Ross (1982) fit very closely to those of Kelly (1982). This analysis indicates that motion processing in the early stages is very fast. At a later point we will give reasons why later stages are probably much slower (see section on the Temporal Integration of Velocity Signals). The model is also consistent with the view that "phi" motion or the motion encoded by the "long range process" which can be seen for delays over 100 msec is not mediated by this mechanism (see Anstis, 1980).

The optimum spacing between input receptive fields for a given velocity can also be estimated by using similar reasoning. The distance can be estimated by noting the visual angle corresponding to a $90°$ phase shift for a given spatial frequency component which is optimal for a given velocity—as such it is one quarter of the spatial period of the most

effective spatial frequency measured at a given velocity. This calculated Δx is plotted as a function of velocity from the empirical curve fits of Kelly's (solid line) and that of Burr and Ross (closed circles) in Fig. 6(A). Note that as velocity is increased in the very slow range, very little change in the optimum spacing occurs but that for faster velocities, the hypothetical spacing becomes proportional to velocity.

Independent measures of ciritical spatial and temporal intervals as a function of velocity have been made and were mentioned earlier (Van Doorn and Koenderink, 1982a,b; Nakayama and Silverman, 1984). These critical intervals are also plotted in Fig. 6(A,B). Despite the very different techniques employed, there appears to be a surprising degree of quantitative agreement between the data obtained from moving sine waves and moving random dots.

Taken together, both sets of results are consistent with the view that many parallel motion mechanisms must operate on a moving image and that different subsets are most responsive for the differing velocity ranges. For very low velocities, it appears that this is handled by somewhat higher spatial frequency mechanisms having a range of different temporal frequency response characteristics. For higher velocities above about 10 deg/sec, different sets of detectors are employed but here they vary mainly in terms of their spatial frequency characteristics.

In addition to the fact that the proposed scheme provides a framework to think about data from random dots and sinusoidal gratings, it also resolves what might be called the 8 Hz paradox, accounting for the otherwise puzzling differences in experimental findings seen for drifting gratings in comparison to stroboscopically illuminated moving objects. We have outlined the fact that optimal temporal frequency to see moving sinewave gratings is about 8 Hz* (Kelly, 1979) and that this same frequency is the best to elicit the MAE and directionally specific adaptation. Yet in many other respects this figure of 8 Hz is a very poor frequency to see motion, especially in situations where real moving targets are stroboscopically illuminated at this rate.

At this 8 Hz strobe rate, observers rate the "quality of motion" as very poor (Sperling, 1976), MAE's are essentially nonexistent (Banks and Kane, 1972), and velocity discrimination deteriorates (McKee and Welch, 1984). In addition, this is the frequency where size constancy (presumably influenced by velocity

---

†This exact figure of 8 Hz should not be taken too literally because of the well-known relationship between temporal sensitivity and mean luminance (van Nes et al., 1967). This variation in luminance is likely to explain the difference in the optimal temporal sensitivity obtained in different studies. For example, Pantle (1974) found a peak sensitivity at about 5 Hz using a mean luminance of 12 cd/m² whereas Burr and Ross (1982) generally found a peak sensitivity at 10 Hz using a mean luminance of 200 cd/m².

mechanisms) is the poorest (Rogowitz, 1983). Finally, this same strobe rate of 8 Hz fails to provide adequate stimulation for the normal development of directional selectivity in kitten cortical neurons (Cynader and Chernenko, 1976). Single cells from cats reared at this strobe rate have normal orientation selectivity but rarely have direction selectivity. These cats are also deficient in direction selectivity when measured behaviorally (Pasternak et al., 1984).

The apparent 8 Hz paradox disappears when we consider our quadrature model. For the continuously drifting sine wave grating, a frequency of 8 Hz stimulates the proposed detectors spaced at a quarter-cycle at a temporal interval of about 33 msec. Thus a continuously drifting sine wave having a temporal frequency of 8 Hz is actually providing the motion system with a sampling rate of 32 Hz rather than 8 Hz. In the case of the 8 Hz strobe illumination of a moving object, however, the stimulation is delivered to adjacent quarter-cycle detecting units much less frequently, at 125 msec intervals. Many other experiments indicate that such a delay between successive frames is inadequate to activate the short range motion process (Braddick, 1974; Pollack, 1972).

Although the present hypothesis has been developed most directly from psychophysical data, it has a number of testable neurophysiological predictions. First it suggests that the direction selectivity of neurons to stroboscopically illuminated moving slits would be optimally stimulated at a rate of approximately 4 times the optimal temporal frequency of the cell when tested with gratings. If we assume an optimal drift rate of 8 Hz, the results are in accord with single unit recordings in cats where it takes strobe rates of 32 Hz for the responses of directionally selective cells in the cortex to behave identically to intermittent as compared to continuously moving stimuli (Cremiux et al., 1984). It also suggests that given a particular spatio-temporal peak frequency sensitivity as measured by drifting gratings, a neuron should have a predictable peak velocity as tested by moving slits. Assuming linearity, the peak velocity sensitivity would be predicted by the optimal spacing divided by the optimal timing as outlined in Fig. 6 [see also equation (2) in Fourier Description of Moving Images].

MOTION HYPERACUITY

Recently there has been widespread interest in visual hyperacuity. Hyperacuity can be defined as a precision of visual localization beyond the resolution limit. The term hyperacuity was introduced in the study of vernier acuity (Westheimer, 1971) and stereoscopic acuity. Hyperacuity thresholds of about 5 arc sec are common, about 1/5 the diameter of the smallest foveal cone. Hyperacuity also occurs in foveal motion sensitivity where a comparable displacement threshold (D_{min}) of 5 arc sec can also be

measured (Nakayama and Tyler. 1981; Nakayama *et al.*, 1984). It also is reported in the periphery (Biederman-Thorson *et al.*, 1971; McKee and Nakayama, 1984) where displacement thresholds are much smaller than two-point or grating resolution.

Hyperacuity begins to lose its apparently paradoxical quality after one reflects on some of the factors which could determine these thresholds. Given an adequate amount of blur before sampling to prevent aliasing and given enough photons, hyperacuity will be limited by neuronal signal-to-noise ratios. In this section we shall not marvel as to why hyperacuity is so good, but will ask why it is so bad.

This point is especially appropriate with regard to motion sensitivity because of its rather large summation area in relation to static position sensitivity (Nakayama and Tyler, 1981). Large amounts of summation can provide the opportunity to increase signal-to-noise ratios. To emphasize this point, consider the striking example of the locust visual system where such summation appears to lower the motion hyperacuity limit several orders of magnitude below the minimum angle of resolution (MAR). Thorson (1966a,b) measured the optomotor response of the desert locust to sinusoidally moving optokinetic drums. First he determined that the highest spatial frequency which would mediate the optomotor response was less than 0.33 c/deg. The locust had a visual acuity which was consistent with its ommatidial spacing, about 200 times worse than the human. Despite this lack of spatial resolution, however, the smallest movement (D_{min}) of the optokinetic drum that elicited a behavioral response from the locust was nearly comparable to the human motion sensitivity, as low as 20 arc sec. Thus the ratio MAR to motion hyperacuity in the locust is about 270:1, more than 50 times the comparable ratio of 5:1 generally seen for the human fovea. So despite the very low spatial resolution of the locust visual system, it has a specialized system for seeing motion which is nearly comparable to our own in terms of displacement thresholds. Presumably this occurs because motion information can be integrated over the whole retina for the optomotor response, with the consequent increase in the signal-to-noise ratio.

Perhaps more directly relevant to human vision are the data on displacement thresholds of retinal primate ganglion cells reported by Scobey and Horowitz (1976). These investigators have found that for peripheral receptive fields having an excitatory width of about 1 deg, the minimum displacement that elicits a threshold response is about 1 arc min. If we equate the minimum angle of resolution with receptive field size, it indicates a MAR to hyperacuity ratio of 60:1. It should be pointed out that psychophysical measurements of the ratio of MAR to motion hyperacuity have been made in the human periphery as well as the fovea and fall very short of this 60:1 value, having a range between 5:1 and 10:1 (see McKee and Nakayama, 1984). This comparison of primate reti-

nal electrophysiology and human psychophysics again emphasizes the point—psychophysically measured motion hyperacuity, rather than showing the exquisite sensitivity of the visual system, actually points to some striking limitations. It suggests that much of the information from peripheral ganglion cells as recorded by Scobey and Horowitz is unavailable to the higher motion centers to assist in making fine motion discriminations. In a study examining the displacement thresholds in sine wave gratings, Nakayama and Silverman (1985) suggest that this limitation is imposed by the existence of an early saturation of the contrast detecting elements which feed into the motion system.

METRICAL ENCODING OF VELOCITY

So far we have concerned ourselves with threshold phenomena: the maximum (D_{max}) or the minimum encodable displacement (D_{min}) or the minimum contrast or signal-to-noise ratio to see motion. These measures tell us little about the encoding of velocity between these two extreme limits, however. Here we deal with the fact that a very useful property of a motion signal is that it is metrical, that it is a measure of velocity, not just an indicator of its presence or direction. Presumably it is this metrical aspect that enables image motion to have such value in doing many of its very useful tasks described earlier.

In psychophysical terms, we can explore metrical precision in terms of the degree to which small changes in the velocity can be detected. In other words, we can define precision as the inverse of the differential velocity threshold. Because velocity is a vector, we need to address two components, magnitude as well as direction.

To determine the precision of the encoding of velocity magnitude, McKee (1981) introduced a paradigm to measure the ability to see differences in velocity for successively presented moving targets. In these studies, the observer was given a large block of trials containing a small range of velocity magnitudes. The task was to say whether a given stimulus was faster or slower than the mean for that block of trials. Except for slow velocities below about 1.5 deg/sec, Weber fractions for velocity were essentially constant, about 5%. The same result obtains in the periphery, but the range over which low velocities have a Weber fraction of greater than 5% now increases in accordance with that expected from increases in the spatial grain of the periphery as estimated from visual acuity measurements (McKee and Nakayama, 1983).

Weber's law for velocity was also found in the study by Nakayama (1981) where he measured an observer's ability to see differential shearing motion in random dots which was accompanied by common image motion. For common motion of greater than 2 arc min, the differential threshold rose in proportion to the common image motion amplitude with a

Weber fraction of $5\%_0$. The adherance to Weber's law in Nakayama's (1981) experiment was maintained at much lower velocities than for the successive presentation as described by McKee (1981). Also consistent with these findings are experiments conducted by Van Doorn and Koenderink (1983) showing an adherence to Weber's Law. Rather than measuring the smallest difference in velocity that could be detected, they measured the signal-to-noise ratio of coherently moving random dots in relation to randomly moving dots for given ratios in velocity between neighboring retinal regions. Although they focused on different quantitative aspects of their results, a close inspection of their data for high signal-to-noise ratios, comparable to that used by Nakayama (1981) and McKee (1981), is also consistent with a Weber fraction of 5%.

The encoding of velocity direction is a separate issue. Using random dots, Levinson and Sekuler (1976) have shown that observers can match velocity directions to about 1 deg. Nakayama and Silverman (1983) have found the same value using complex line stimuli.

If we consider image velocity to be represented in a two-dimensional velocity space with the horizontal and vertical axes representing horizontal and vertical components of velocity, respectively, then it is possible to outline a hypothetical set of two-dimensional figures which represent the boundaries of velocity discrimination for given velocities. Given the value of 5% for magnitude, a value of 1.0 deg for direction, and a Cartesian "velocity space", discriminability ellipses for velocity can be envisioned. The results to date indicate that these ellipses have a major:minor axis ratio exceeding 3:1 (see Fig. 7).

FOURIER DOMAIN DESCRIPTION OF MOVING IMAGES

Visual image movement is a spatial/temporal event and it can be represented mathematically in terms of a luminance function $L(x, y, t)$ of space and time. Alternatively, the same image motion can be equivalently expressed in the frequency domain where $L(Fx, Fy, Ft)$ is the Fourier transform of $L(x, y, t)$. Because it appears that the early processing of visual information involves a simultaneous and rather efficient encoding of both position and spatial frequency (Campbell et al., 1969; DeValois et al., 1982; Marcelja, 1980), it is also useful to think of image motion in spatio-temporal frequency terms.

Thanks to recent papers (Fahle and Poggio, 1981; Kelly, 1982; Watson and Ahumada, 1983), some important features of moving stimuli in the frequency domain can be outlined. Considering only the horizontal spatial dimension, Fig. 8 represents image movement in spatio-temporal frequency terms. Points A and A' represent the loci of spectral energy of a spatial sine wave drifting to the right and points B and B' represent the loci for a sine wave of the same
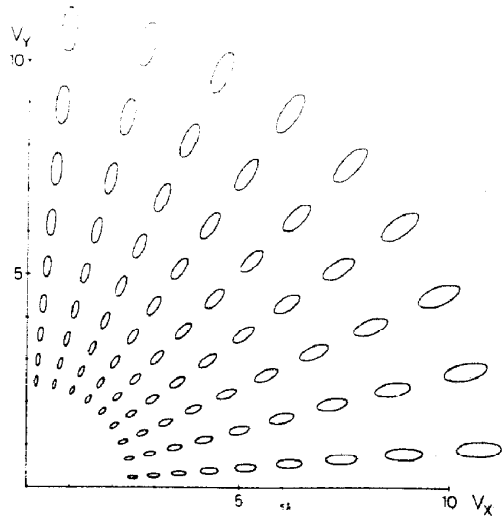

Fig. 7. Estimated discriminability ellipses for motion. One quadrant of a hypothetical four quadrant "velocity space" is represented where V_x and V_y designate horizontal and vertical velocities, respectively. Scale is deg/sec.

spatial frequency drifting to the left at the same velocity. If they have the same amplitude and are summed, the result is a counterphase grating (Levinson and Sekuler, 1975). The velocity of any drifting sinusoidal component is simply the temporal fre-
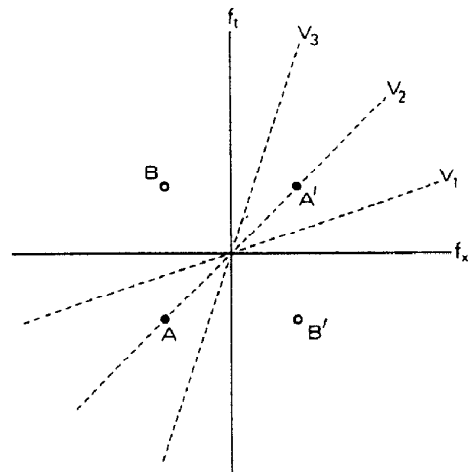

Fig. 8. Spatio-temporal Fourier representation of visual image motion in the horizontal direction. Abscissa (f_x) represents the horizontal spatial frequency axis, ordinate (f_t) represents the temporal frequency axis. The vertical spatial frequency axis cannot be shown in this two-dimensional representation. Two dots labeled A and A' represent the locus of spectral energy of a pure sinusoidal grating moving to the left. Dots labeled B and B' represent the locus of spectral energy of a similar grating moving to the right. If the contrast of these two gratings is the same, the four dots represent a counterphase grating. Dashed line labeled V_2 represents the locus of spectral energy of all possible sinusoidal components moving to the left with the velocity of the leftward moving grating. This is a line passing through the origin conforming to the equation $V_x = TF/SF$. Lines labeled V_1 and V_3 are combinations of spatio-temporal frequency energy which move with velocities one third and three times the velocity of V_2, respectively.

quency divided by the spatial frequency:

$$V_r = \text{TF/SF} \qquad (2)$$

thus any constant ratio of temporal to spatial frequency defines a particular velocity and this is represented by a straight line through the origin (see sets of dotted lines which depict lines of different velocities).

As a consequence of equation (2) any coherent image motion, regardless of its spatial frequency composition, has spatio-temporal frequency energy restricted to a single line passing through the origin. In the more general two-dimensional spatial case, coherently moving optical stimuli have amplitude components confined to a single plane passing through the origin of the corresponding three-dimensional Fourier space (Kronauer et al., 1983). Furthermore the tangent of the dihedral angle that this plane forms with the Fx, Fy plane defines the velocity magnitude.

Apparent motion provides a specific example of the power of a Fourier domain representation of moving stimuli (Watson and Ahumada, 1983). A vertical line moving continuously in the horizontal direction has a spatial-temporal Fourier transform described by a line passing through the origin (see lines V1, V2, V3 in Fig. 8). Watson and Ahumada (1983) note that if this motion is sampled intermittently (at instants spaced apart by the temporal interval Δt), the spatio-temporal Fourier spectrum will be elaborated with spectral energy also occurring on parallel "replicas" (Fig. 9). The spacing of these replicas becomes closer as the sampling interval Δt increases (see Bracewell, 1965; Morgan, 1980). Of primary interest to Watson and Ahumada was the boundary condition where apparent motion was psychophysically indistinguishable from real motion. They made the conjecture that it occurs when the parallel "replicas" are outside the window of visibility as defined by the highest spatial and temporal frequencies that can be picked up by our visual system. Furthermore they made some empirical observations which provided an estimate of these boundaries and show that they were in rough accord with previous psychophysical estimates.

It should be noted that sampled motion will also contain spectral energy components corresponding to movement in the opposite direction. Ordinarily an observer does not see such reversed motion even though linear theory would predict that channels sensitive to this portion of the spatio-temporal energy spectrum would be stimulated.* As such, these oppositely moving components may be masked or suppressed. In a later section, we shall see that frequency domain characteristics of moving stimuli may also help to clarify additional transformation of signals in the motion system.

---

*A preliminary description of the characteristics of such channels can be seen in papers by Holub et al. (1981) and Thompson (1984).
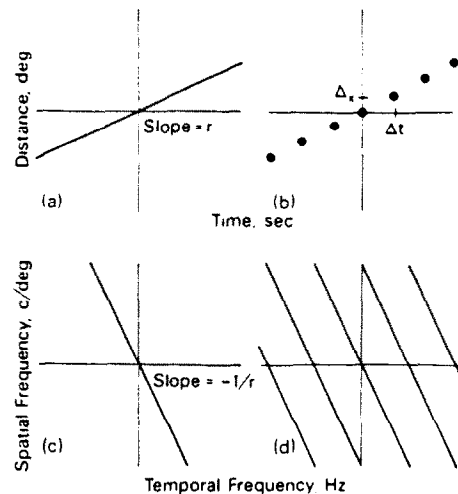
Fig. 9. Real and apparent motion represented in space–time (top) and in terms of spatial and temporal frequency (bottom). From Watson et al. (1983).

CHROMATIC INPUT TO THE MOTION SYSTEM?

Research to date indicates that color may provide little if any input to early motion processing. A provocative experiment in this regard was conducted by Ramachandran and Gregory (1977) who replicated Braddick's (1974) original "short range" experiment using random pixels which were either red or green. As they adjusted the luminance balance between the red and the green and approached the point of isoluminance, the emergent figure, which was previously segregated, disappeared. To answer criticism that the stimulus did not have sufficient chromatic contrast due to much lower spatial frequency sensitivity of the red-green system (Van der Horst and Bouman, 1969), they also used very large pixels (1 deg). Since such a pattern has high spectral power in the low spatial frequency region, it is unlikely that the lack of motion detection was due to lack of a chromatic contrast signal.

More recently, Cavanagh et al. (1984) measured perceived velocity in chromatic and luminance modulated red-green sinusoidal gratings. As the luminance of the red bars was brought very close to that of the green bars, a dramatic slowing of the perceived velocity of the grating was seen and this was most prominent for the lowest spatial frequencies, it cannot be explained away by reduced high spatial frequency sensitivity characteristics of the chromatic system. As such it means that relative to the luminance contribution to the motion system, the chromatic input must be very weak.

COMPUTATIONAL THEORIES OF MOTION PROCESSING

Early models

Some of the most important early studies of motion processing were accompanied by an algorithmic

theory. In particular, the earliest explicit model was proposed by Reichardt (1961), formulated to account for the characteristics of the insect optomotor response. In the most simple terms, the basic theoretical unit of motion detection was a pair of receptors such that the delayed outputs of one receptor were multiplied by the output of the other. Partly as a consequence of the very strong nonlinear property of multiplication, the theoretical motion signal had a number of peculiar features which were in accordance with the data obtained from insects. First it predicted a square law relationship between luminance and the motion signal. Second it predicted the existence of a reversed motion response in two situations. Reversed motion would be seen for gratings which had half periods smaller than the interommatidial distance (showing the property of spatial aliasing) and it would also be seen for the special case where stepping motion was accompanied by a reversal of contrast at each step (see also Anstis, 1970; Anstis and Rogers, 1975). Finally it predicted a range of spatial and temporal frequences (and hence velocities) that would enable the mechanism to yield an appropriate directional response.

The next model for motion processing was neurophysiological. Barlow and Levick (1965) suggested two theoretically possible mechanisms to mediate directional selectivity in rabbit retinal ganglion cells: a system having unidirectional lateral excitation in the preferred direction and a system having unidirectional lateral inhibition in the nonpreferred direction. In the first case a moving spot would fall on parts of the retina which were facilitated by its earlier presence in a neighboring region and in the second case the spot would fall on retina that would be inhibited by the earlier presence of the spot. Either or both properties would endow the system with primitive directional selectivity. To evaluate these possibilities, Barlow and Levick (1965) used a two slit experiment. They presented a second slit at a variable distance and a variable time after the first slit, both in the preferred and nonpreferred direction of the cell.

The results were decisive. They demonstrated that it was inhibition that played the major role, vetoing the responses to movement in the nonpreferred direction. The inhibition had a certain rise time and decay. If stimuli were moved faster in the nonpreferred direction so as to arrive at a second site before the inhibition, the directional selectivity was lost. This was also the case for very slow movements in the nonpreferred direction, where the lateral asymmetric inhibition had been given the chance to decay. Thus Barlow and Levick's motion system had the property of being directional over a specific range of velocities. The relative importance of inhibition in mediating a cell's directional selectivity has been confirmed in many different structures and animals using the same Barlow and Levick paradigm (Ganz and Felder, 1984; Michael, 1965). The role of inhibition also

received independent support from experiments where inhibitory synaptic transmission was disrupted pharmacologically with a consequent abolition of directional selectivity (Wyatt and Daw, 1976; Sillito, 1977).

In a quantitative follow-up, Emerson and Coleman (1981) determined that the movement of a stimulus through a cortical cell's receptive field produces a response which corresponds very closely to the linear summation of individual flashes to each separate subregion of the traverse, doing so only in the preferred direction. In the nonpreferred direction they noted a nonlinearity consistent with the inhibition already proposed by Barlow and Levick (1965). To account for such nonlinearities mechanistically, a specific cellular mechanism of shunting inhibition as opposed to subtractive inhibition has been proposed (Thorson, 1964; Torre and Poggio, 1980). This could be mediated at the same postsynaptic membrane (Ariel and Daw, 1982; Baylor and Fettiplace, 1979).

Despite the major differences between Reichardts computational model and Barlow and Levicks's neural model, there are some important similarities. The neurophysiological model is consistent with some aspects of the "multiplier" idea insofar as the vetoing inhibition furnishes at least part of the multiplicative process (see Thorson, 1966a). Another point of similarity is that each model specifies direction but no metrical value of velocity magnitude.

A third early model of motion processing comes not from biology, but from engineering—arising from the practical need to measure the optical speed of moving objects for specialized applications. A number of authors addressed the problem of providing a noncontact measure of the optical speed, for example the speed of a rolled steel plate as it passes by, or the speed of the ground from an aircraft taking aerial photographs (Ator, 1963, 1966; Agar and Blythe, 1968). The basic component is a parallel slit reticle which occludes a single photosensor. Ideally, the device is very narrowly tuned to just one spatial frequency in the image. Assuming that the target of interest contains a wide range of spatial frequencies, including that of the reticle, the accurate calculation of velocity can be made by dividing the temporal frequency of the photosensor output by the spatial frequency of the reticle [see equation (2)]. An important difference between this model and others is that it measures optical speed. It does not measure direction of motion.

A fourth computational model makes a simultaneous measurement of the spatial and temporal gradients of illumination along the x and y axes (Hadani et al., 1980; Horn and Schunk, 1981). As mentioned earlier, this model [as embodied in equation (1)] is unlikely to mediate mammalian motion sensitivity.

Recent models

Since the formulation of these four classes of

motion models, it has become increasingly apparent that at least in the mammalian visual system, the early processing of visual information consists of channels responding to different spatial frequencies or different spatial scales of the image (Marr, 1982). It is perhaps the recognition of this single development that best differentiates early from later models.

A recent model of van Santen and Sperling (1984) capitalizes on this view by proposing a modified Reichardt model which has as its front end such band-limited spatial frequency channels. One major advantage of the model is that it eliminates the spatial aliasing of the original Reichardt model and thereby does not predict the existence of a reversed motion percept to any continuously moving object which is dominated by spatial frequencies whose half period is smaller than the inter-detector spacing. This lack of spatial aliasing for continuous motion is consistent with human motion perception. van Santen and Sperling (1984) also varied the contrast between adjacent panels and were able to verify the predicted existence of the multiplicative relationship over a small range of contrast as predicted by the model.

The paper by van Santen and Sperling also provides an analysis of Reichardt type models in general, showing that despite the existence of a highly non-linear multiplier stage, there are some surprising quasilinear characteristics (also see Thorson, 1966a). In particular, different temporal frequency components are shown to superpose linearly. Accordingly, the model makes the prediction that for stimuli containing more than one spatial sinusoidal component, a constant velocity of all components is not the optimal stimulus. Optimal is the somewhat unusual stimulus where each spatial frequency component slides past the other at its optimal temporal frequency.

Marr and Ullman (1981) use spatial frequency filtering to provide a more plausible realization of the spatio-temporal gradient model mentioned earlier. They suggest that the image is convolved with a receptive field operator (del^2G, Laplacian of a Gaussian). This convolution generates a new image intensity distribution I_c, where the same logic embodied in equation (1) applies. Thus:

$$V_x = \frac{\partial \nabla^2 G * I / \partial x}{\partial \nabla^2 G * I / \partial t}. \qquad (1a)$$

To reduce the number of computations, measurements are restricted to those at the zero crossings of I_c, corresponding to regions containing significant changes in image luminance. For an even greater reduction of the computational load, the algorithm simplifies the operation implied by equation (1a) by considering only the sign of the numerator and denominator. Velocity information is lost, but the algorithm calculates direction very economically, requiring only the very primitive comparison of two signed signals.

Marr and Ullman suggest that Y cells perform the required operation of taking the time derivative. Such cells, however, are fed by rectifying subunits (Hochstein and Shapley, 1976) and are unlikely to transmit the sign of $\partial \nabla^2 G * I / \partial t$. Although this speaks against Marr and Ullman's hypothesized role for Y cells, the computational aspects of the model stands as a new theoretical contribution.

All models described so far have highly nonlinear components at an early stage. A moving stimulus, however, would seem to be adequately described in terms of linear components. In the spatio-temporal frequency domain, for example, a moving grating is simply described by two spectral dots placed symmetrically about the origin (refer back to Fig. 8). Thus the moving grating could be picked up by a linear channel tuned to a spatio-temporal band of spectral energy which brackets these dots. This appears to be the reasoning adopted by Watson and Ahumada (1983) in their presentation of a hypothetical linear motion sensor. In their model, a pair of spatially tuned receptive fields are arranged in spatial and temporal quadrature for all spatial and temporal frequencies and the outputs of these two parallel channels are simply added, not multiplied. A somewhat more physiologically plausible linear model was proposed by Adelson and Bergen (1984). van Santen and Sperling (1984) argue that both of these so-called linear models, require the introduction of non-linearities at later stages and become formally equivalent to the Reichardt model.

It should be noted however that none of these simple models has any feature which would give a metrical readout of velocity. Being linear sensors, they simply reflect the amount of spectral energy integrated by the particular channel. The outputs of such channels could be used in comparison with the outputs of many other linear channels, however, to provide a metrical reading of motion. A similar type of between-channel comparison would also be necessary to obtain velocity magnitude information from a multichannel version of the van Santen and Sperling model as well (1983).

A hypothetical arrangement to accomplish this between-channel comparison is to have higher-order units summate the outputs of spatio-temporal filters which share a common velocity. As such they would be arranged along a radial line of constant velocity (see Fig. 8). Velocity could be read out by comparing activity in these different higher order radial "velocity" channels. This could be determined by detecting the mode or the peak of the population profile response possibly with the aid of lateral inhibition. Preliminary evidence suggests that this type of neuronal organization may occur in MT, an extrastriate area specialized for image motion (Newsome et al., 1983).

Alternatively, if a very different principle of temporal frequency encoding were used (see Ator, 1963; Watson and Ahumada, 1984), a comparison across

"velocity" channels might be unnecessary. In this case velocity magnitude could be read out directly by bypassing the time averaging suggested in the linear models and simply reading the raw temporal frequency (TF) from the linear channel and scaling it to read angular velocity by dividing by the preferred spatial frequency of the channel. Koenderink (1983) noted that the receptive fields which are extremely narrowly tuned with respect to spatial frequency (Pollen and Ronner, 1975) could provide the most accurate readout of velocity if configured according to the design proposed by Ator (1963). This model assumes that temporal frequency can be read by the visual system. As yet there is no neurophysiological evidence for temporal frequency encoding in the visual system, but the idea has some precedent in the auditory system where it would seem that the fine structure of the temporal impulse rate carries information regarding frequency (Kiang, 1965). A potential difficulty with this temporal frequency readout model, however, is that precise velocity encoding can occur for durations and temporal frequencies that expose the system to just a fraction of a temporal cycle (McKee *et al.*, 1984). Furthermore in experiments comparing the discrimination of velocity at threshold, it appears that the underlying mechanisms to detect temporal frequency must be very broadly tuned (Thompson, 1984).

A second approach which could bypass a comparison between velocity channels would be to reinstate metrical precision into the Marr/Ullman model so that it computes velocity in accordance with equation (1a).

Beyond the simple pair?

Most of the above models treat the motion detecting apparatus as a spatio-temporal pair of detectors. As such, it would suggest that a pulse pair consisting of a stimulus in one position followed by a stimulus in an adjacent position would constitute both an adequate and optimal stimulus for the motion system. Evidence from some rather diverse paradigms, however, suggests that this pulse pair stimulus may not be the optimal stimulus for the motion system and that motion detectors are sampling spatio-temporal energy over more than two spatio-temporal positions.

One of the earliest investigators to emphasize this point was Sperling (1976) who noted that qualitatively, two spatio-temporal pulses were far inferior in eliciting strong "goodness of motion" reports from observers, especially in comparison to a longer sequence of multiple spatio-temporal pulses. Although this observation was neglected for some time, it has been the recent focus of a number of independent studies using a variety of paradigms. Lappin and Fuqua (1982) found that increasing the number of different spatio-temporal frames greatly increased the probability of detection of motion in random dot displays. McKee and Welch (1985) found that dis-

crimination of timing using sequential pulse pairs was better than that expected from the simple probability summation of individual pulse pairs. In a Braddick type paradigm, Nakayama and Silverman (1984) found that when two successive displacements of random dots were made within about a 50–200 msec interval, the D_{max} per displacement could rise by $50\%_{0}$.

A Fourier representation of the two pulse motion stimulus is consistent with this view. Rather than being restricted to a single line through the origin as is the case for a continuously moving object (Fig. 8), the spatio-temporal Fourier amplitude spectrum will be nonzero almost everywhere. If we consider the spatio-temporal Fourier plane as in Fig. 8 and regard amplitude as the height above and below this plane, the Fourier amplitude spectrum is a corrugated "cosine" surface with the locus of one of its peaks being a straight line passing over the origin. One would expect that a heterogeneous set of early sensors encoding a wide range of velocities would be activated by this distribution of spectral energy, increasing the ambiguity of any motion signal.

Both the empirical results and the Fourier domain considerations suggest that most of the models presented so far are probably incomplete in a fundamental sense and that future motion models will require a set of input detectors that includes more than a single pair.

SINGLE CELL ANALYSIS OF IMAGE MOTION

Definitional issues

Single unit recordings in a variety of visual structures reveal an overwhelming number of cells which respond vigorously to moving stimuli. In fact their number is embarrassingly large unless one thinks that most of the visual brain is exclusively devoted to the analysis of motion. Given the number of visual functions related to the pick-up of information in moving images and perhaps not related to encoding moving objects, the number is more reasonable. It should also be recognized that a moving stimulus traverses many retinal points and has greater opportunity to stimulate cells which may also be responsive to stationary targets. In any event, this widespread activation of visual neurons by moving stimuli raises a definitional question. What do we mean by a motion sensitive cell and how does this cell differ from other visual cells?

Some of the most relevant papers appeared very early. In the process of characterizing several new classes of cell in the rabbit retina, Barlow *et al.* (1964) were careful to distinguish true direction selectivity from what they termed *misleading direction selectivity* (italics mine). In particular they noted that rabbit cells similar to Kuffler (1952) units could masquerade as directionally selective especially when the antagonistic summation zones were not quite circularly symmetric. This apparent directional selectivity reversed its sign if the contrast of the test target was

reversed (also see Hubel and Wiesel, 1962, 1965; Albus, 1980). Thus the apparent direction selectivity of many visual cells could be explained by linear summation of the responses over various portions of the receptive field. For example, in the rabbit retina it could arise from the coincidence of excitation associated with a spot entering an ON zone with the excitation rebounding with the removal of inhibition associated with the spot leaving an OFF zone. Barlow and Hill (1963) demonstrated that for their truly directionally selective cells, their response was essentially invariant over a wide range of contrasts and most critically, directional selectivity remained the same for stimuli having reversed contrast.

A set of criteria for definining directional selectivity in single neurons is implicit in the experimental procedures described by Barlow et al. (1964). In particular, a movement selective neuron should alter its firing rate only for changes in velocity and direction and not for other parameters. That is, one could say that a cell is coding motion if, and only if, it shows the property of orthogonality; its response can be altered only by variations in the velocity of the stimulus and not by any other stimulus dimension at least over a significant physiological range. This definition of a channel in terms of orthogonality has been proposed by Regan (1982).

This is a reasonable concept, yet closer inspection of many of the neurophysiological results since the original discoveries of Barlow et al. (1964) suggest that adherance to this definition may be too restrictive. It is suggested that important aspects of velocity encoding would be missed if one defined motion sensitivity so narrowly. Consider a quantitative study of the response properties of certain movement neurons in the frog optic tectum (Grüsser and Grüsser-Cornehls, 1973). They report that within a significant stimulus range, the impulse rate of the neuron increases monotonically as a power function of velocity having an exponent of 0.7, clearly a strong dependence on velocity. The neuron was also sensitive to contrast as well as area, however, and a more complete description of the firing rate of the neuron was described as

$$R = k v^{0.7} C^{0.55} \log (A/A^*) \qquad (3)$$

where R is impulse rate, V is velocity, C is contrast, A is area, and A^* is a threshold area.

Describing the response properties of such neurons in terms of distinct categories or "trigger features" seems premature. Velocity is a major determinant of the firing rate but it is equally clear that other aspects of the stimulus contribute a significant share in influencing the discharge rate. A second example can be seen in mammalian striate cortex where the discharge rate of directionally selective units can be influenced by contrast, temporal frequency, and spatial frequency as well as velocity (Holub and Morton, 1981; Albrecht and Hamilton, 1982). Common to the two cases is the fact that each of the neurons is influenced by variations in stimulus velocity and that the property of orthogonality is clearly violated.

At this point it should be stressed that although orthogonality is a convenient mathematical property of coordinate systems, any linearly independent set of nonorthogonal basis vectors is sufficient to represent a locus in some multidimensional space. Linear transformations can convert any nonorthogonal set of basis vectors into an orthogonal set if required.

But the ever-present problem of noise in biological systems lends some small advantage for detection schemes more closely orthogonal. The vestibular coding of the three dimensions of head angular velocity is a familiar example. Each pair of semicircular canals defines a basis vector, oriented perpendicular to the plane of a canal. To encode all three dimensions of head rotation requires that all basis vectors not lie in a plane (that they are not linearly dependent). To obtain an optimum signal/noise ratio, however, requires that these basis vectors should also be orthogonal (see Robinson, 1983). The advantage of strict orthogonality is small, such that moderately large departures (up to 30 deg) will have negligible effects in the efficiency of encoding.

We now return to the problem of recovering velocity information from neurons whose response varies with a number of stimulus dimensions. The frog tectal unit described in equation (3) could provide unambiguous velocity information if it were combined with information from at least two other units which were less sensitive to velocity, one more sensitive to contrast and the other to size. Formally, this might entail the solution of the required number of simultaneous equations to disambiguate each of these sensory dimensions (Richards et al., 1982).

To summarize this digression on orthogonality, it is helpful to encode the primary dimensions orthogonally but this is not decisive because moderate departures from orthogonality do not prevent the recovery of information. This means that if the discharge rate of a neuron is covarying to some considerable degree with the stimulus velocity or direction, it could carry velocity information with essentially the same fidelity as an orthogonal "labeled line" or channel.

This does not mean that orthogonality is not attained at some level of the visual system or the brain. Our point is that orthogonality with respect to velocity could be arrived at in successive stages and that to ignore classes of visual neurons that lack orthogonality may obscure an understanding of how motion is encoded in successive stages.

For this reason we suggest the designation of some cells as "pre-movement" sensitive. Such cells would provide the input to the motion system but could also have the capability of mediating other visual functions. Hypothetical cells with responses described in equation (3) for example, qualify for such a role because their response will covary with the velocity of

the stimulus, yet they will not be motion sensitive if strictly defined in terms of orthogonality. By adopting the term "pre-movement" we also imply that some early cells in the visual pathway may not be pre-movement.

Nonmovement, pre-movement and movement units

At this point, we examine neuronal classes at several different levels of the visual pathway, considering their possible relation to image motion processing. First we can ask whether all classes of photoreceptors are pre-movement. To the extent that chromatic input plays little if any role in the processing of motion (see earlier), it is likely that the blue cone photoreceptors are not pre-movement. This follows from data which suggests that the blue cones make no contribution to luminance but only to color (Eisner and MacLeod, 1981).

We can then ask the same question for other known cell classes at different levels. For example does motion processing derive its input from particular sub-classes of bipolar, ganglion, LGN and striate cortical cells? More pointedly, are there significant generic classes of early elements which are outside the pathway of motion processing. We ask this question because clear answers might simplify the problem, reducing motion processing to its essential elements.

Consider the circularly symmetric receptive fields of primates with concentric ON and OFF regions. These are found in the retina, lateral geniculate nucleus (LGN) and in the early input stage of primate striate cortex. None of these cells in the retina or LGN are directionally selective, yet it must be the conjoint activity of some of these cells that conveys the needed information to higher order neurons which are directionally selective. We ask whether particular classes of these concentric RFs provide input to motion processing. Primate LGN neurons fall into six distinct cellular layers. Neurons in the four most dorsal layers, the parvocellular laminae, are very sensitive to chromatic differences and have low sensitivity to luminance modulation. They make up the majority of X cells in the LGN. The remaining neurons are those located in the magnocellular layers of the LGN. These receive a broad band chromatic input without color opponency and have a high contrast sensitivity (Kaplan and Shapley, 1982).

Isoluminant chromatic stimuli can lead to very reduced sensations of motion (Cavanagh et al., 1984) and to an abolition of the figure–ground segregation based on motion (Ramachandran and Gregory, 1978). This casts doubt as to whether the very large number of chromatic neurons of the parvocellular LGN can provide significant input to early stages of motion sensitivity. This view is further strengthened if we consider the possibility of "spurious" stimulation for individual achromatic cells even at the point of psychophysically defined isoluminance. The exact weighting of different cone inputs is likely to vary from cell to cell with consequent variation in the isoluminant point of each cell. So an isoluminant grating for one achromatic cell would not be isoluminant for another. Thus the psychophysical criterion of isoluminance does not eliminate the stimulation of achromatic neurons. Consequently, the existence of residual motion at the point of isoluminance could be mediated by achromatic cells.

It is recognized that the opponent cells in the parvocellular layers do have some achromatic contrast sensitivity, but their contrast thresholds are very high (Kaplan and Shapley, 1982; Hicks et al., 1983; Lennie et al., 1984), about an order of magnitude higher than cells in the magnocellular layers, at least for medium and low spatial frequencies. The elevated contrast thresholds of these cells are too high to have a significant role in mediating the observed contrast sensitivity functions measured using motion direction as a criterion (Burr and Ross, 1982).

The contrast independence of two other motion related phenomena constitute independent pieces of evidence. Adding additional contrast beyond about 4–5 times the contrast threshold does not increase motion aftereffect strength or directionally specific adaptation (Keck et al., 1976; Keck et al., 1980; Pantle et al., 1978; Sekuler et al., 1978). Increasing the contrast above 2–3% does not lower motion thresholds (D_{min}) in sinusoidal gratings (Nakayama and Silverman, 1985).

Because parvocellular cells code only at the high end of the dynamic range of contrast, all of these results support the simplifying conjecture that these cells have no significant input to the early stages of the motion system. It would appear that a most significant portion of motion sensitivity is mediated by the cells in the magnocellular layers of the LGN. These cells can be subdivided into two classes, X and Y (Blakemore and Vital-Durand, 1981; Kaplan and Shapley, 1982).

Y cells comprise a class of neurons which have been suggested as mediating motion sensitivity (Tolhurst, 1973), mainly on the basis that they respond preferentially to transient stimuli. It could be argued, however, that these cells are very unlikely to mediate early motion sensitivity, at least the type used for the ordinary coding of continuous real motion, because Y cells do not code spatial phase (Hochstein and Shapley, 1976). We suggest that spatial phase (or position) is one of the fundamental building blocks upon which motion processing must rest as motion is a change in phase over time. To develop this point, first consider the possible role of X ganglion cells in mediating motion sensitivity.

Information regarding motion direction can be recovered from X cells if one has additional information regarding the polarity of the target contrast. Consider an ON center X cell. Suppose we had a bright edge at the border between center and surround and moved it slightly closer to the receptive field center. The cell would fire more. Just the op-

posite would happen if we moved a dark edge closer to the RF center. Assuming the existence of other systems, which could sense the contrast polarity of the local edge, information as to the X cell's increment or decrement in firing rate could provide information as to direction. A somewhat more elaborate version of this argument was originally outlined by Marr and Ullman (1981).

An experimental way to assess the importance of X as opposed to Y cell input for motion might be to devise a situation where Y cells but not X cells are activated sequentially across the retina; then to use psychophysical criteria to determine whether the short range motion process is activated. Lelkins and Koenderink (1984) set up a field of random dots which was mostly stationary and unchanging, but one panel in this display differed from the other regions insofar as the dots were continuously replaced by other random dots. To provide the opportunity to see "illusory movement" Lelkins and Koenderink had this region of dynamic perturbation move in a single direction. It should be clear that there is no linear spatio-temporal energy moving along the path of the "disturbance" and as a consequence one should not expect there to be an associated increase in average X cell population activity which moves along with this disturbance. Furthermore, none of the computational models described earlier would respond to this motion. The "disturbance" itself is indeed moving in a particular direction and the psychophysical experiments indicate that human observers did report some movement. But Lelkins and Koenderink found that the perception of movement for this particular pattern was weak and essentially different from that ordinarily elicited by real motion. It did not lead to the classical motion aftereffect nor did it have the capacity to drive optokinetic nystagmus. From the nonlinear properties of Y cell, it is expected that this moving disturbance would provide strong Y cell activation across the retina in register with the disturbance (Hochstein and Shapley, 1976). Because the stimulus does not lead to the ancillary phenomena ordinarily associated with the "short range" motion process, in particular by not eliciting a motion aftereffect, it provides evidence that the sequential activation of Y cells is not involved in the encoding of real motion.

If the above reasoning is correct, it raises the possibility that only a relatively small number of cells in the LGN (comprising about 15% of all the cells) are "pre-movement", namely the X cells of the magnocellular layers. Furthermore, it is known that such cells have well defined target sublaminae of striate cortex, terminating in layer IV C alpha (Blasdel and Lund, 1983) and that a major projection from this lamina goes (via layer IVB) to MT, an extrastriate area specialized for motion.

Motion processing at the extrastriate level

Extrastriate areas have also been examined and much attention is focused on MT, a rather small cortical projection field on the posterior bank of the superior temporal sulcus. MT receives major projections from striate cortex (V1) and (V2). Furthermore, much of the afferent input appears to be myelinated, suggesting the specialization of this area for rapid processing of visual information (Van Essen, 1979).

Single unit recordings in MT of the primate show marked directional selectivity in comparison to striate cortex and other extrastriate areas, suggesting a special role for this area in the processing of motion (Zeki, 1974a, 1974b; Maunsell and Van Essen, 1983a; Albright *et al.*, 1984). Recent behavioral experiments involving lesions of MT also support its presumptive role in mediating motion sensitivity. By making a very small restricted chemical lesion in MT, also sparing fibers of passage in the nearby optic radiations, Newsome *et al.* (1983) were able to show a specific oculomotor deficit in the matching of smooth pursuit eye velocity to target velocity. This is consistent with the view that the velocity signals required for pursuit (Rashbass, 1961) are also mediated by MT, a view also consistent with major anatomical projections from MT to specific vermal areas of the cerebellum via the dorsolateral pontine nuclei. Both of these areas are implicated in the mediation of smooth eye movements and show modulated discharge to target velocity and eye velocity (Suzuki *et al.*, 1981; Suzuki and Keller, 1984).

Additional evidence to link MT with motion processing comes from two very recent findings by Movshon and colleagues, also providing important insight as to how visual signals undergo functional transformation in the motion system. The first finding is that some MT cells respond best to the true direction of a compound set of gratings having differing orientations rather than to its underlying spatial Fourier components (Movshon *et al.*, 1984). These results are discussed in greater detail as part of the section on Orientation Tuning.

Second is the increasing covariance of the response with respect to the stimulus velocity at the expense of reduced covariance with respect to spatial and temporal frequency. This point can be appreciated by comparing the spatio-temporal frequency response properties of single neurons in striate cortex with those recorded from MT. Striate cortical neurons are responsive to a restricted spatio-temporal region of frequency space as demonstrated by recordings in cat cortex (Holub and Morton, 1981), having two-dimensional spatio-temporal frequency tuning functions which can be roughly decomposed into the product of separate spatial and temporal frequency responses.

Combinations of spatio-temporal frequency energy which have constant velocity, however, lie along a 45° line in a log frequency representation. Thus a two-dimensional response profile of a cell picking up a constant velocity could not be defined as the simple

product of temporal and spatial frequency sensitivity. Of interest is the fact that some cells in MT have their maximum sensitivity oriented along these constant velocity lines and thus appear to represent a higher order process, abstracting velocity despite rather major variations of spatial and temporal frequency (Newsome et al., 1983).

Extrastriate analysis of image motion beyond MT has not been examined in detail, but scattered observations suggest additional processing of motion information. First is the presence of major anatomical projections of MT to other extrastriate cortical areas in the parietal-occipital area, specifically to area MST and to the posterior parietal area 7a (Maunsell and Van Essen, 1983b). Second is the clear specialization for image motion for some neurons in these areas. Mountcastle and Motter (1981) found units sensitive to motion in area 7a, with many responding to a radial expansion or contraction of the image.

ORIENTATION TUNING IN THE MOTION SYSTEM

Are the receptive fields of the units feeding directionally selective neurons tuned to orientation? For motion selective systems in sub-cortical visual centers (tectum, accessory optic system, etc.), there is little evidence for orientation selective neurons in these or antecedent structures. As such, motion sensitivity is thought to be mediated by spatial input filters having circular symmetry.

In the case of the mammalian visual cortex, it is not immediately clear as to the nature of the input receptive fields which feed motion sensitive mechanisms. Many cortical cells are orientation selective and direction selective which might suggest that the input filters to motion cells are themselves oriented, but this conclusion is not compelling. Cortical cells with elongated receptive fields, for example, could be made up of a linear array of local motion detectors, each having a concentric input receptive field. Furthermore, the existence of circularly symmetric cells in the layer IV of primate striate cortex as well as the existence of large numbers of circularly symmetric RFs in supragranular layers of cortex having high levels of metabolic activity (Livingston and Hubel, 1984) indicates that the input to cortical motion cells need not be orientation selective. Dow (1974), for example, found movement sensitive cells in layer IVb of money striate cortex which are poorly tuned with respect to orientation. It is recognized that cells in this lamina project to MT. As we have previously noted, this extrastriate area appears specialized for image motion.

Some research seems to support the idea of a broader angular tuning for motion than orientation, which might suggest that motion systems are fed by input stages having little or no orientation selectivity. Ball and Sekuler (1979) used coherently moving random dots and added masking motion in many simultaneous directions. They found that masking was greatest by moving dot displays having motion components in the same direction as the test direction. The angular tuning width of this masking function was very broad, much broader than the tuning widths of cortical neurons tuned to orientation (DeValois et al., 1982) or psychophysical estimates of orientation selectivity (Campbell and Kulikowski, 1968; Blakemore and Nachmias, 1971).

Related studies have also been conducted in single cortical neurons. Hammond (1981) tested complex cells of the striate cortex for directional selectivity using two types of pattern: oriented lines moving in a direction orthogonal to their orientation and isotropic random texture. They found that the directional tuning functions corresponding to each stimulus could be very different. In many instances, the isotropic random dots had a much broader directional tuning function than the lines.

From these findings one might conclude that motion and orientation are separately encoded. Yet the surprising difference in directional tuning for different stimuli within the same cell (Hammond, 1981) was troubling, raising the specter of additional complexity in an already complicated picture of neural processing in cat area 17. To clarify this problem, Movshon et al. (1980) made an important point regarding the three-dimensional Fourier representation of moving textures composed of isotropic random dots. These stimuli contain a whole set of spatio-temporal frequency components oriented very far from the true direction of motion. In fact they are spread over a plane in Fx, Fy, Ft space which passes through the origin (see section on Fourier Description of Moving Images). This means that moving spatial frequency components span a range of ± 90 deg from the direction of motion. Consequently random dot stimuli are not sufficiently selective to answer questions regarding the relationship between orientation and direction selectivity.

Gizzi et al. (1981) also argued that other extrastriate cortical systems such as the lateral suprasylvian area (LS) of the cat are also highly tuned to orientation even though initial work using spots suggested otherwise (Spear and Baumann, 1975). Employing sine wave gratings, they were able to show a rather high degree of orientation tuning in LS. Again, the apparently discrepant conclusion lies in the fact that a single moving spot contains a wide range of component orientations and velocities.

The importance of orientation tuning in motion processing receives independent support from psychophysical experiments which suggest an elongated receptive field for motion. Nakayama et al. (1985) compared differential compression and shearing thresholds in random dots. As mentioned earlier, Nakayama and Tyler (1981) have shown that thresholds for differential shearing motion rise rather rapidly above about 0.7 c/deg. Nakayama et al. (1985) extended this observation to include compression

thresholds. They demonstrated that the rapid rise in threshold for higher movement spatial frequencies seen for shearing motion is not nearly so pronounced for compression motion. These results are consistent with the notion of an elongated receptive field for motion where the direction of motion is orthogonal to its major axis. Such a receptive field can be estimated to have a length of about 15 arc min and a width of 5 arc min in central vision.

An oblique effect for motion?

The issue of orientation tuning leads to a consideration of the "oblique effect", the superior contrast sensitivity and grating resolution for vertical and horizontal gratings as compared to obliquely oriented gratings (Freeman et al., 1966; Campbell and Kulikowski, 1968). Most efforts to measure motion sensitivity as a function of orientation have not shown an oblique effect. This raises the question as to whether the mechanisms used to detect static contrast are the same as those which feed into the motion system. In a systematic study, Ball and Sekuler (1980) found no differences in sensitivity for any direction of motion using random dot stimuli. This study employed 3 separate measures of motion sensitivity: reaction time, contrast thresholds, and motion aftereffect duration. As mentioned earlier, however, moving random dot stimuli contain spatial frequency components spanning a wide range of orientations and correspondingly, the range of the directions of all components for a given direction of motion of random dots spans 180°. As such it is unlikely that such stimuli isolate mechanisms sensitive to particular orientations so it is not surprising that no oblique effect was found using random dots. Pasternak and Merigan (1980) replicated the lack of an oblique effect for motion in man and cat, and also extended it by making some measurements using low frequency square wave gratings.

One additional reason why motion appears to be isotropic is that the static oblique effect is mainly confined to higher spatial frequencies (Campbell et al., 1966). To the extent that directionally selective mechanisms get their inputs from RFs having larger receptive field sizes, an oblique effect would not be expected.

The aperture problem

We assume that motion is encoded by motion-selective units which derive their inputs from localized receptive fields. This is essentially equivalent to viewing a moving object through an aperture. Suppose an extended line moves through this aperture [see Fig. 10(A)]. The velocity in the aperture can be described by V_L, the local velocity orthogonal to the orientation of the line. This local velocity is insufficient to specify the true direction of the moving line because V_L could be generated by an infinite set of true velocity vectors V [see Fig. 10(B)]. Thus an analysis of local motion cannot specify the true velocity to anything better than 180 deg. This is often referred to as the aperture problem and has received extended discussion (Marr, 1982; Marr and Ullman, 1981; see also Wallach, 1935).

Because of this 180° ambiguity, it would seem that a single local reading of velocity is not very informative. Marr and Ullman (1981) suggest that many local readings would be necessary to narrow down the range of possibilities. If one thinks in terms of the *magnitude* and the direction of the local reading of velocity, however, it can be shown that a single reading is highly informative since it constrains the true target velocity to fall along a straight line defined by the equation $V = V_L/\cos(\theta)$, where V is the magnitude of the true velocity vector, θ its possible direction, and V_L the local reading of velocity. This is represented by the dotted line in Fig. 10(B) (see Fennema and Thompson, 1979; Adelson and Movshon, 1982).

It follows that the existence of a real moving object having two orientations can provide sufficient information for a pair of orientation-selective velocity detectors to reconstruct the real direction and velocity (Fennema and Thompson, 1979; Adelson and
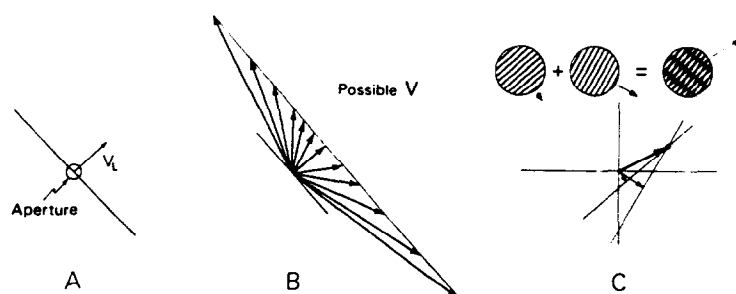
Fig. 10. Pictoral description of the aperture problem. Consider a local window designated by the circular region. In (A) the local velocity V_L is orthogonal to the contour and is moving up and right at a 45° angle. In (B), are shown the possible real motion vectors V which could have given rise to V_L, this defines a constraint line (indicated by the dashed line). In (C), we show one situation elucidated by Adelson and Movshon (1981) where two gratings moving at different velocities have intersecting constraint lines which identify a compound motion which is clearly not the vector summation of independent component motions.

Movshon, 1981). More are superfluous for the case where noise is not a factor and where the movement can be assumed to be a rigid two-dimensional translation of the image (see Hildreth, 1983). The solution proposed is not the linear vector sum of the component directions but is formally equivalent to finding the intersection of constraint lines in "velocity" space. From Fig. 10(C) it should be obvious that this intersection is not the vector sum of the two local components. Adelson and Movshon (1981) measured perceived direction with compound gratings and found that the perceived direction was indeed in qualitative accordance with the solution in velocity space and this has been confirmed more quantitatively by Daugman (1981).

Before providing an extended discussion on further implications of Adelson and Movshon's original experiment, we deal with some alternative formulations. First is the issue of how the compound grating might be represented in three-dimensional spatial-temporal frequency space. Such a moving compound grating forms two pairs of dots symmetric about the origin and it follows that these spectral components define a plane passing through the origin whose orientation is consistent with the perceived direction and velocity of the compound grating stimulus (refer to the section on Fourier Domain Description of Moving Images). As such, the velocity of the compound grating can be determined by finding the plane in STF space, defined by these pairs of points in the 3-D spectrum. Formally, the solution in 3-D frequency space is a mathematically more general case of finding the point in velocity space as suggested by Adelson and Movshon (1981). As such, velocity and spatio-temporal frequency representations are consistent.

A more serious question to be posed is whether one needs to postulate a higher stage of motion at all. Perhaps the visual system can simply track the nodes of the compound grating. Suppose for example that there were an early nonlinearity in the contrast response function. Mixtures of sinusoidal gratings would then generate intermodulation distortion products which are of sufficient amplitude to carry the motion percept using only very low level receptive fields. If this were the case, it might be argued that no synthesis of differently oriented velocity signals would be required. A similar point was made by Daugman (1981) who argued that complex gratings contain a "missing fundamental", one which could be the result of nonlinear processing. Although Daugman's point applies for cases where the temporal frequency of each of the component gratings is equal, it does not generalize to the case where the temporal frequencies are different. In this latter case the direction of movement of a compound grating as predicted by the velocity constraint model is no longer the same as that predicted by a simple model based on nonlinear distortion products. An intuitive way to understand this problem is to realize that even if
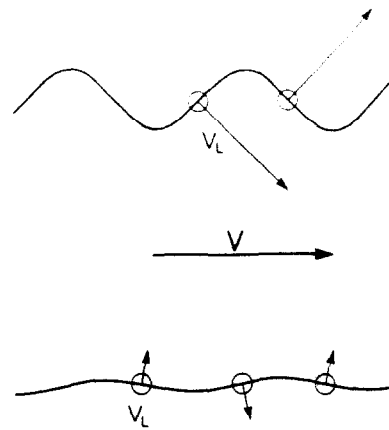

Fig. 11. Planar sine wave moves to the right with velocity V. Note that the largest local velocities are much smaller than V and also in very different direction. If the maximum angular difference in the sine wave is low, each individual local motion component moves with perceptible independence and the pattern is seen as moving nonrigidly (from Nakayama and Silverman, 1983).

nonlinear distortion products were formed, they would also have the form of moving gratings and one would be faced with yet a new aperture problem.

More simply, one can also conceive of premovement sensitive elements which are linear, having a local Kuffler-like concentric receptive field organization. Although, there is no spatio-temporal frequency energy in the direction of the moving nodes, the Fourier interpretation of the stimulus is somewhat misleading because a population of X cells having linear concentric receptive fields will be sequentially activated along the path of these moving nodes. Such a model is not inconsistent with various subcortical motion systems that do not show any orientation tuning.

An experimental argument against these alternative ideas comes from a 2 × 2 adaptation paradigm (Movshon et al., 1984) where test and adapting stimuli were horizontally moving gratings or horizontally moving diagonal plaid patterns consisting of 2 obliquely oriented gratings. Adapting to plaids had a selective effect on elevating direction specific thresholds to plaids and vertical gratings selectively elevated direction-selective thresholds to vertical test gratings. Such results are difficult to interpret in terms of distortion products or localized motion adaptation which is fed by linear circularly symmetric receptive fields and support the role of orientation tuning.

In real life another situation is perhaps as common, namely, that the differently oriented contours are in different spatial positions. Take the case of sine waves moving across the page with velocity V (see Fig. 11). Note that there are no oriented components moving with the real velocity V, and the absolute value of the velocity of any of the components is much lower than the real velocity. Therefore, if the movement of this object is to be correctly encoded from information furnished by oriented velocity units, information

needs to be combined from orientation units from separate regions of the visual field rather than from within the same region. This could be accomplished using the same reasoning proposed by Adelson and Movshon (outlined in Figure 10C) extending it slightly so as to include the interaction of velocity signals from different retinal regions.

This simple theory, however, needs some modification for their are obvious cases when the aperture problem is not solved. For example, moving sine waves making an angle of less than about 15° at their zero crossing, appear as obviously nonrigid (Nakayama and Silverman, 1983). An example of such a perceptually nonrigid stimulus is seen in the bottom portion of Fig. 11. Here the local components are very far from the real motion vector and to solve for the real velocity requires one to find the locus of intersection for constraint lines which cross at very shallow angles. Lines intersecting at shallow angles are inherently difficult to localize if there is any uncertainty as to the orientation or position. Given the likelihood of noise associated with these oriented velocity signals, one might postulate that the visual system defaults to an interpretation where the different directions of the local motion are simply accepted. Consequently, the observer sees an illusory nonrigidity.

So far we have only considered the special case where the true velocity vector is the same all over the two-dimensional image. In many situations this is not the case, objects can rotate in the plane, outside the plane, and the existence of several separate objects can lead to a heterogeneous amount of true image motion on the retina. Likewise for the case of an object undergoing a rigid deformation.

The fact that such motions are possible raises some rather unexpected theoretical difficulties for the encoding of the velocity field, an important insight derived from the computational perspective (Ullman, 1983). Take the case outlined by Hildreth (1983). Figure 12 shows a succession of two frames before and after a compound motion, consisting of a translation, rotation and deformation. The velocity vector at any point P is unspecified. One heuristic strategy to estimate velocity at all points along the curve is to
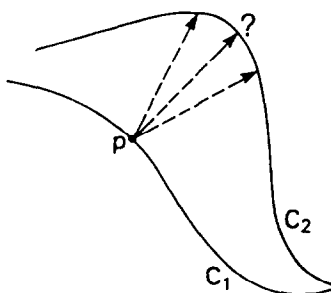

Fig. 12. If Curve C_1 rotates, translates and deforms over time to yield curve C_2, the velocity of the point p is ambiguous (from Hildreth, 1983).

assume that the local velocity vectors vary smoothly over neighboring regions of the visual field (see Horn and Schunk, 1981). One of the problems with this approach is that real image velocity vectors can be discontinuous at the edges of real objects, thus a two-dimensional smoothness constraint will obscure the pick-up of biologically significant information. A mathematical algorithm proposed by Hildreth (1983) may reduce this problem by restricting the smoothness constraint to one dimension, along contours rather than across two directions across the visual field. Her algorithm incorporating this one-dimensional smoothness constraint generated a set of theoretical velocity fields which were in qualitative accord with a large number of perceptual illusions associated with two-dimensional figures undergoing rigid rotation. It does not account for the nonrigidity seen for some figures undergoing pure translation, however (see Fig. 11).

TEMPORAL INTEGRATION OF VELOCITY SIGNALS

Earlier [see Fig. 3(A)] we hypothesized the existence of a distinct stage where velocity signals are temporally integrated, noting that such an integrator was an essential component of several computational models of motion processing (Reichardt, 1961; Foster, 1971; van Santen and Sperling, 1984). Integration has the important property of smoothing out local phase sensitive fluctuations from early detecting stages and increases the effective signal-to-noise ratio.

Empirical support for such an integrator comes from insects and humans. Srinivasan (1983) recorded from directionally selective interneurons of the fly visual system. Such cells are responsive over very large regions of the visual field and are thought to mediate the optomotor response. Srinivasan found a reciprocity for short duration high velocity "pulses" of a moving grating and longer duration slower velocity "pulses". Each produced the same monophasically exponentially decaying response with a time-velocity reciprocity occurring up to about 60 msec. A quantitative analysis of the gain and phase of the locust optomoter response is also supportive of the integrator concept. Thorson (1966a) found that the gain fell off rapidly above 0.5 Hz with a phase lag increasing above this frequency. This very limited temporal resolution of the insect motion system is very different from earlier elements which show no roll-off for frequencies as high as 50 Hz (French, 1980).

A number of human psychophysical observations are also supportive of the integrator concept. Velocity thresholds for sinusoidal shearing motion in random dots shows extreme low pass characteristics (see Fig. 13) much different than the results obtained for homogeneous flicker sensitivity (de Lange, 1952). The slope of this low pass curve is compatible with two first order integrators having a time constant of about
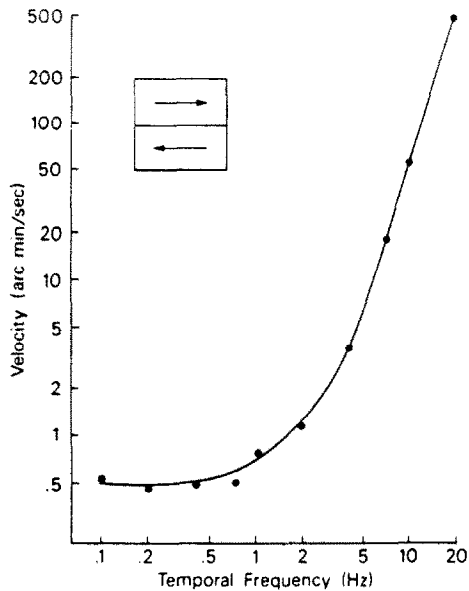
Fig. 13. Temporal frequency sensitivity of motion sensitivity. The stimulus is differential shearing motion in random dots, oscillated sinusoidally. Threshold velocity is plotted against temporal frequency. Note the relatively low frequency characteristics of the system with sensitivity falling off at about 1 Hz (from Nakayama, 1985).

80 msec (see Nakayama and Tyler, 1981; Nakayama, 1984). Another indication of velocity integration is the fact that the velocity signals from successive pairs of displacements show nonlinear facilitation which can persist for 200–400 msec (Nakayama and Silverman, 1984). McKee (1984) has found that suprathreshold judgments of velocity magnitude are more precise when each moving stimulus is well separated in time from the other. When two different velocities were closely spaced in time, the observer had great difficulty in perceiving velocity differences. Finally, Regan and Beverley (1984) have demonstrated that the minimum velocity threshold continues to decrease up to duration of 500 msec.

At this point we make distinction between this hypothesized velocity integrator and the existence of a very different velocity integrator, one specialized to drive optokinetic nystagmus (Collewijn, 1982). The optokinetic gain only reaches its full asymptotic value after a very long period of between 10–20 sec, suggesting an integrator with a very long time-constant, on the order of 20 sec. This is further supported by the existence of optokinetic after-nystagmus (OKAN) which consists of slow eye movements in the same direction as the original OKN and which lasts for many seconds if an observer is kept in darkness. From the time-course of the build up of OKN and the decay of OKAN, it should be clear that this must be a very different integrator than the one proposed here. At present it is thought to complement the much faster acting and somewhat highpass temporal frequency characteristics of the vestibulo-ocular reflex (Henn, 1979).

HIGHER-ORDER COMPUTATIONS ON THE OPTICAL FLOW FIELD

So far we have dealt with the basic constituent elements of motion processing. Here we describe how this motion information may be further organized, combining motion information from one portion of the visual field with another or combining motion information from the two eyes.

Derivatives of velocity

Gibson (1950) was one of the first to emphasize the importance of the optical velocity field as a source of information regarding the layout of three-dimensional space. He also suggested that the raw velocity field itself was not as informative as the gradient of the velocity field because raw velocities might vary greatly with changes in observer direction or speed, whereas the gradient might not.

The gradient of the optical velocity field is more complex than gradients usually encountered in textbook treatments of vector analysis because the gradient operator is customarily applied to a scalar field to define a vector field. With the optical velocity field, we already have a vector field, and taking a gradient of this two-dimensional vector leads to a 2×2 matrix containing four spatial first derivatives of velocity (dV_x/dx, dV_x/dy, dV_y/dx, dV_y/dx) (see Koenderink and Van Doorn, 1976; Longuet-Higgins and Prazdny, 1980). It can be shown that Gibson was technically incorrect in stating that the gradient remained invariant with changes in observer motion but Koenderink and Van Doorn (1976) have demonstrated that one component of the gradient is invariant with surface layout.

For an observer moving with respect to a planar surface, Koenderink and Van Doorn (1976) specified a generalized first-order expansion of the optical flow field, defining three components of the gradient: curl, dilation, and deformation

$$\text{Grad } V = \text{Curl} + \text{Div} + \text{Def}$$

and indicated that Def remained invariant under a wide variety of changes in observer motion, whilst Curl and Div did not. Longuet-Higgins and Prazdny (1980) approached the issue in reverse, asking how an observer can determine the local slant of an arbitrary surface using information about higher-order spatial as well as temporal derivatives of the velocity field.

In theory, these linear differential operators contain useful biological information. An issue is whether they are literally computed by the visual system. In the previous sections it was established that the motion signal has lowpass spatial and temporal frequency characteristics. This indicates that the motion signal must be integrated both temporally and spatially. Therefore the hypothetical derivative or differencing operators are probably very coarse, because they combine velocity information from large receptive fields and over rather long temporal intervals.

In addition to this bandwidth limitation, the hypothetical computation of these derivatives of the optical velocity field is complicated by the Weber law for velocity (Nakayama, 1981; Van Doorn and Koenderink, 1982). Taking a derivative implies taking a difference, but the ratio metric underlying Weber's law complicates the reconstruction of a signal based on differences. It is cumbersome to reconstruct a signal proportional to any spatial derivative in the face of common image motion insofar as the taking of derivatives implies taking a difference rather than a ratio. This suggests that theories which imply the extraction of useful information from spatial derivatives of the velocity field (Nakayama and Loomis, 1974; Koenderink and Van Doorn, 1976; Longuet-Higgins and Prazdny, 1980) should not be taken too literally.

Of the three derivative operators, divergence has received the most experimental attention. Regan and Beverley (1978, 1983) suggested that a neural network sensitive to changing size, mathematically equivalent to computing a divergence, could be a useful adjunct in the perception of motion in depth. Regan and Beverley suggested that there are changing size detectors in the visual pathway, specialized channels that could be adapted with prolonged exposure. Their experimental strategy was to adapt to changing size by oscillating the sides of a square. As a control, they moved the square sideways along a diagonal such that the space-time average velocity of each individual contour was the same for both kinds of motion. Under a wide variety of conditions they found that one could obtain increased threshold elevations for changing size as compared to detecting sideways motion. They concluded that there are specialized detectors in the visual pathway for changing size.

To show that such detectors are indeed responsive to changing size, other detecting schemes must be ruled out. A "convexity" detector was suggested by Nakayama and Loomis (1974). This hypothetical neural unit sums excitatory velocity information from a central region and combines it with inhibiting signals from a concentric surround. It generalizes these differences for all orientations of movement. Such units would be sensitive to both divergence and curl and would become adapted to the stimuli utilized by Regan and Beverley. If such units were underlying the results of Regan and Beverley's looming detection, one would predict that moving stimuli having divergence would also show cross-adaptation to curl.

An independent and somewhat theory-based counter-argument against a curl detector can also be raised. Julesz and Hesse (1970) generated patterns which varied in curl in different portions of the display. It consisted of line elements (needles) which had random orientations. All of the needles in the majority of the display rotated with the same velocity. In a smaller section of the display, a group of needles rotated in the opposite direction. To the extent that

curl is a basic sensory feature, a *texton* in Julesz's terminology (1984), the two areas should appear as perceptually segregated. No segregation was seen. This indicates that either the criterion of pre-attentive segregation is not an appropriate one to identify basic sensory dimensions or that curl is not such a primary dimension.

Interocular comparison of velocity signals: motion in depth

Thanks to the work of Regan and colleagues, it appears that there is a well-defined set of motion analyzers with specific connections between the two eyes. Beverley and Regan (1973) adapted human observers to targets moving sinusoidally along different horizontal lines in three-dimensional space. Then they tested threshold sensitivity to detect the same range of 3-D motions and evaluated the relationship between test and adapting directions. They isolated 4 distinct "motion-in-depth" channels. Two channels were responsive to motion in depth along axes which essentially "missed" the head; one channel specialized for the trajectory of motion which passed to the left of the left eye, the other for motion which passed to the right of the right eye. The other two channels are specialized for motion between the two eyes, one for motion to the right of the midline, the other to the left. In a separate paper, Beverley and Regan (1975) suggested a Hering-type of opponent process model for the discrimination of direction. They confirmed this view by measuring the directional discrimination of motion in 3D. Sensitivity was highest just where the differences in sensitivity between the channels was most steeply varying. As such, the discrimination results were in good agreement with the results obtained with adaptation (Beverley and Regan, 1975).

One can think of at least two possibilities as to the makeup of these channels. First these channels could be formed by the binocular combination of monocular velocity signals alone. Second, the motion-in-depth signal could arise after the encoding of binocular disparity. This would require yet another separate motion system having as its primitives, differing positions in a disparity and direction space. Apart from the fact that this second alternative seems needlessly complicated, additional evidence against its existence can be cited. Richards and Regan (1973) made perimetric measurements comparing motion-in-depth sensitivity with static stereoacuity. Large variations in each dimension were found over different regions of the visual field and the variations were often uncorrelated. This supports a separate mechanism for motion-in-depth and for static stereopsis. To form these motion-in-depth channels the system needs to link the specific polarity of motion direction in the two eyes and to also compare their absolute magnitudes with some considerable precision. In the two outside channels the direction of motion is the same direction for the two eyes, and the

difference in magnitude determines whether the motion is moving on a trajectory passing to the left or the right of the head. For the two inside channels signifying a motion trajectory which is between the two eyes, the polarity of the motion in each eye is different. The absolute magnitude determines whether it is left of or right of the midline.

Parallel single unit recordings from the parastriate cortex of cat provide neurophysiological support for the analysis of motion-in-depth mediated by monocular motion sensitivity. Early work (Pettigrew, 1973; Zeki, 1974a) suggested the existence of neurons preferentially sensitive to opposing motion in the two eyes. More systematic work by Cynader and Regan (1978, 1982) shows the existence of motion-in-depth channels which respond over a very wide range of binocular disparities. Some neurons were preferentially tuned to different directions of motion in three-dimensional space and represent a separate system from that used for the coding of static binocular disparity.

CONCLUDING REMARKS

Biological visual systems have specialized mechanisms to detect the movement of optical images. Although much remains to be discovered, it appears that the pick-up of motion information is beneficial for a wide variety of visual tasks. This includes a role in reconstructing the third dimension, segmenting the image, driving eye movements, eliciting attention, encoding self motion, mediating size constancy, and detecting moving objects.

In this review, we have emphasized the results of human psychophysics and the recordings of single neurons in the geniculostriate system in primates. We have extrapolated beyond the data, speculating on the broad features of this hypothetical geniculostriate motion system. Future work should recognize the need for explanation at many levels as suggested by Marr (1982). Thus simultaneous efforts in ecological optics, computational theory, psychophysics, neurophysiology and neuroanatomy are of potential importance in providing a satisfying picture of image motion processing.

In addition to the importance of explanation at different levels, there is the likelihood that more than one motion system exists and that a separate multilevel analysis of each might be useful. Future studies using techniques of stimulus manipulation but measuring something other than simple motion detection are likely to determine the degree to which various motion systems participate in different visual funcitons. This could include an examination of the stimulus characteristics underlying vection, OKN, image segmentation, pursuit eye movements, depth reconstruction, etc. It is possible that a small number of motion systems will be more clearly delineated, allowing a structural, functional and computational description of each.

REFERENCES

Adams R. (1834) An account of a peculiar optical phenomenon seen after having looked at a moving body, etc. *London and Edinb. Phil. Mag. J. Science.* 3rd series, **5**, 373–374.

Adelson E. H. and Bergen J. R. (1985) Spatiotemporal energy models for the perception of motion. *J. opt. Soc. Am. A* **2**, 284–299.

Adelson E. H. and Movshon J. A. (1982) Phenomenal coherence of moving visual patterns. *Nature* **300**, 523–525.

Agar W. O. and Blythe J. H. (1983) An optical method of measuring transverse surface velocity. *J. scient. Instrum., J. Phys. E,* 1968, Series 2, **1**, 25–28.

Albrecht D. G. and Hamilton D. B. (1982) Striate cortex of monkey and cat: contrast response function. *J. Neurophysiol.* **48**, 217–237.

Albright T. D., Desimone R. and Gross C. G. (1984) Columnar organization of directionally selective cells in visual area MT of the macaque. *J. Neurophysiol.* **51**, 16–31.

Albus K. (1980) The detection of movement direction and effects of contrast reversal in the cat's striate cortex. *Vision Res.* **20**, 289–293.

Allman J. M., Miezin F. and McGuinness E. (1984) Direction and velocity specific surrounds in three cortical visual areas of the owl monkey. *Perception.*

Anstis S. (1970) Phi movement as a subtraction process. *Vision Res.* **10**, 1411–1430.

Anstis S. M. (1978) Apparent movement. In *Perception,* Vol VIII, *Handbook of Sensory Physiology.*

Anstis S. M. (1980) The perception of apparent movement. *Phil. Trans. R. Soc. Lond. B* **290**, 153–168.

Anstis S. and Rogers B. J. (1975) Illusory reversal of visual depth and movement during change of contrast. *Vision Res.* **15**, 957–962.

Ariel M. and Daw N. W. (1982) Pharmacological analysis of directionally sensitive rabbit retinal ganglion cells. *J. Physiol.* **324**, 161–185.

Atkinson J. (1979) Development of optokinetic nystagmus in the human infant and monkey infant: an analogue to development in kittens. In *Developmental Neurobiology of Vision* (Edited by Freeman R. D.). Plenum Press, New York.

Atkinson J. and Braddick O. (1981) Development of optokinetic nystagmus in infants: an indicator of cortical binocularity. In *Eye Movements: Cognition and Visual Perception* (Edited by Fisher D. F., Monty R. A. and Sarden J. W.). Erlbaum, Hillsdale, NJ.

Ator J. T. (1963) Image velocity sensing with parallel-slit reticles. *J. opt. Soc. Am.* **53**, 1416–1422.

Ator J. T. (1966) Image velocity sensing by optical correlation. *Appl. Opt.* **5**, 1325–1331.

Attneave F. (1954) Some informational aspects of visual perception. *Psychol. Rev.* **61**, 183–193.

Baker C. L. and Braddick O. J. (1982) The basis of area and dot number effects in random dot motion perception. *Vision Res.* **22**, 1253–1260.

Baker C. L. and Braddick O. J. (1982) Does segregation of differently moving areas depend on relative or absolute displacement? *Vision Res.* **22**, 851–856.

Ball K. and Sekuler R. (1979) Masking of motion by broad band and filtered directional noise. *Percept. Psychophys.* **26**, 206–214.

Ball K. and Sekuler R. (1980) Human vision favors centrifugal motion. *Perception* **9**, 317–325.

Ball K. and Sekuler R. (1980) Visibility of motion in presence of broad band and filtered directional noise. *Perception*.

Banks W. F. and Kane D. A. (1972) Discontinuity of seen motion reduces the visual motion aftereffect. *Percept. Psychophys.* **12**, 69–72.

Barlow H. B. (1953) Summation and inhibition in the frog's retina. *J. Physiol.* **119**, 69–88.

Barlow H. B. and Hill R. M. (1963) Selective sensitivity to direction of motion in ganglion cells in the rabbits retina. *Science* **139**, 412–414.

Barlow H. B. and Levick W. R. (1965) The mechanism of directionally selective units in the rabbit's retina. *J. Physiol.* **178**, 477–504.

Barlow H. B., Hill R. M. and Levick W. R. (1964) Retinal ganglion cells responding selectively to direction and speed of image motion in the rabbit. *J. Physiol.* **173**, 377–407.

Baylor D. A. and Fettiplace R. (1979) Synaptic drive and impulse generation in ganglion cells of turtle retina. *J. Physiol.* **288**, 107–127.

Beverley K. I. and Regan D. (1973) Evidence for the existence of neural mechanisms selectively sensitive to the direction of movement in space. *J. Physiol.* **235**, 17–29.

Beverley K. I. and Regan D. (1975) The relation between discrimination and sensitivity in the perception of motion in depth. *J. Physiol.* **249**, 387–398.

Biederman-Thorson M., Thorson J. and Lange G. D. (1971) Apparent movement due to closely spaced sequentially flashed dots in the human peripheral field of vision. *Vision Res.* **11**, 889–903.

Blakemore C. and Campbell F. W. (1969) On the existence of neurons in the human visual system selectively sensitive to the orientation and size of retinal images. *J. Physiol., Lond.* **203**, 237–260.

Blakemore C. and Nachmias J. (1971) The orientation specificity of two visual aftereffects. *J. Physiol.* **213**, 157–174.

Blasdel G. G. and Lund J. (1983) Termination of afferent axons in macaque striate cortex. *J. Neurosci.* **3**, 1389–1413.

Bracewell R. (1965) *The Fourier Transform and Its Applications.* McGraw Hill, New York.

Braddick O. (1974) A short-range process in apparent motion. *Vision Res.* **14**, 519–527.

Braddick O. J. (1980) Low level and high level processes in apparent motion. *Phil. Trans. R. Soc. Lond. B* **290**, 137–151.

Braitenberg V. and Taddei-Ferretti C. (1966) Landing reaction of *Musca domestica. Naturwissenschaften* **53**, 155–156.

Brandt Th., Dichgans J. and Koenig E. (1973) Differential effects of central versus peripheral vision on egocentric and exocentric motion perception. *Expl Brain Res.* **16**, 476–491.

Bridgeman B. (1972) Visual receptive fields sensitive to absolute and relative motion during tracking. *Science* **178**, 1106–1108.

Burr D. C. and Ross J. (1982) Contrast sensitivity at high velocities *Vision Res.* **22**, 479–484.

Burt P. (1984) Multi-resolution flow through motion analysis. Technical Report, RCA Laboratories, Princeton, NJ.

Campbell F. W. and Kulikowski J. J. (1966) Orientation selectivity of the human visual system. *J. Physiol.* **187**, 437–445.

Campbell F. W., Kulikowski J. J. and Levinson E. (1966) The effect of orientation on the visual resolution of gratings. *J. Physiol.* **187**, 427–436.

Campbell F. W., Cooper G. F. and Enroth-Cugell C. (1969) The spatial selectivity of visual cortical neurons to moving gratings. *J. Physiol.* **203**, 223–235.

Cavanagh P., Tyler C. W. and Favreau O. (1984) Perceived velocity of moving chromatic gratings. *J. opt. Soc. Am. A* **1**, 893–899.

Chang J. J. and Julesz B. (1983) Displacement limits for spatial frequency filtered random-dot cinematograms in apparent motion. *Vison Res.* **23**, 1379–1385.

Coggshall J. C. (1972) The landing response and visual processing in the milkweed bug, *Oncopeltus fasciatus. J. exp. Biol.* **52**, 401–414.

Collett T. (1972) Visual neurons in the anterior optic tract of the privet hawk moth. *J. comp. Physiol.* **78**, 396–433.

Collewijn H. (1982) *The Optokinetic System in Models of Oculomotor Behavior and Control* (Edited by Zuber B. L.). CRC Press, Boca Raton, FL.

Cremieux J., Orban G. and Duysens J. (1984) Responses of cat visual cortical cells to continuously and stroboscopically illuminated moving light slits compared. *Vision Res.* **24**, 449–458.

Cynader M. and Chernenko G. (1976) Abolition of direction selectivity in the visual cortex of the cat. *Science* **193**, 504–505.

Cynader M. and Regan D. (1978) Neurons in cat parastriate cortex sensitive to the direction of motion in three-dimensional space. *J. Physiol.* **274**, 549–569.

Cynader M. and Regan D. (1982) Neurons in cat visual cortex tuned to the direction of motion in depth: effect of positional disparity. *Vision Res.* **22**, 967–982.

Daugman J. (1981) The visual spatio-temporal fundamental. *J. opt. Soc. Am.* **71**, 1632.

DeValois R. L., Yund E. W. and Hepler N. (1982) The orientation and direction selectivity of cells in macaque visual cortex. *Vision Res.* **22**, 531–544.

Dichgans J. and Brandt T. (1978) Visual-vestibular interaction: effects on self motion perception and postural control. In *Handbook of Sensory Physiology* (Edited by Held R., Leibowitz H. and Teuber H. L.), pp. 755–804.

Dimmick R. C. and Karl J. C. (1930) The effect of exposure time upon the Reiz Limen of visible motion. *J. exp. Psychol.* **13**, 365–369.

Dow B. M. (1974) Functional classes of cells and their laminar distribution in monkey visual cortex. *J. Neurophysiol.* **37**, 927–946.

Eisner A. and MacLeod D. I. A. (1980) Blue sensitive cones do not contribute to luminance. *J. opt. Soc. Am.* **70**, 121–123.

Emerson R. C. and Coleman L. (1981) Does image movement have a special nature for neurons in the cat's striate cortex? *Invest. Ophthal. visual Sci.* **20**, 766–783.

Enroth-Cugell C. and Robson J. (1966) The contrast sensitivity of retinal ganglion cells in the cat. *J. Physiol.* **187**, 512–552.

Fahle M. and Poggio T. (1981) Visual hyperacuity: spatio-temporal interpolation in human vision. *Proc. R. Soc. Lond. B* **213**, 415–477.

Fennema C. L. and Thompson W. B. (1979) Velocity determination in scenes containing several moving images. *Comput. Graphics Image Process* **9**, 301–315.

Foster D. H. (1971) A model of the human visual system in its response to certain classes of moving stimuli. *Kybernetik* **8**, 69–84.

Freeman R. D., Mitchell D. E. and Westheimer G. (1967) Effect of orientation on the modulation sensitivity for interference fringes on the retina. *J. opt Soc. Am.* **57**, 246–249.

French A. S. (1980) Phototransduction in the fly compound

eye exhibits temporal resonances and a pure time delay. *Nature* **283**, 200–202.

Frost B. J. (1978) Moving background patterns after directionally specific responses of pigeon tectal neurons. *Brain Res.* **151**, 599–603.

Frost B. J. and Nakayama K. (1983) Single visual neurons code opposing motion independent of direction. *Science* **220**, 744–745.

Frost B. J., Scilley P. L. and Wong S. C. P. (1981) Moving background patterns reveal double opponency of directionally specific pigeon tectal neurons. *Expl Brain Res.* **43**, 173–185.

Ganz L. and Felder L. (1984) Mechanism of directional selectivity in simple neurons of cat's visual cortex analyzed with stationary flash sequences. *J. Neurophysiol.* **51**, 294–324.

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

Gibson J. J. (1954) Visually controlled locomotion and visual orientation in animals. *Br. J. Psychol.*

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

Gibson J. J. (1968) What gives rise to the perception of motion. *Psychol. Rev.* **75**, 335–346.

Gizzi M. S., Katz E. and Movshon J. A. (1981) Orientation selectivity in the cat's lateral suprasylvian visual cortex. *Invest. Ophthal. visual Sci., Suppl.* **20**, 149.

Gogel W. C. (1980) The sensing of retinal motion. *Percept. Psychophys.* **28**, 155–163.

Gogel W. C. (1982) Analysis of the perception of motion concomitant with lateral motion of the head. *Percept. Psychophys.* **32**, 241–250.

Gogel W. C. and Tietz J. D. (1973) Absolute motion parallax and the specific distance tendency. *Percept. Psychophys.* **13**, 284–292.

Golumb B., Anderson R., Nakayama K., MacLeod D. I. A. and Wong A. (1985) Thresholds for shearing motion detection in monkey and man. *Vision Res.* In press.

Goodman L. J. (1960) The landing response of insects. I. the landing response of the fly *Lucilia sericata* and other celliphorinae. *J. exp. Biol.* **37**, 854–878.

Gotz K. G. (1968) Flight control in *Drosophilia* by visual perception of motion. *Kybernetik* **4**, 199–208.

Gotz K. G. (1972) Processing of cues from the moving environment in the *Drosophilia* navigation system. In *Information Processing in the Nervous System of Anthropods* (Edited by Wehnee R.), pp. 255–263. Springer, Berlin.

Graham C. H., Baker K. E., Hecht M. and Lloyd V. V. (1948) Factors influencing thresholds for monocular movement parallax. *J. exp. Psychol.* **38**, 205–223.

Graham M. and Rogers B. (1982) Simultaneous and successive contrast effects in the perception of depth from motion-parallax and stereoscopic information. *Perception* **11**, 247–262.

Grasse K., Cynader M. S., Douglas R. M. (1984) Alterations in response properties in the lateral and dorsal nuclei of the cat accessory optic system following visual cortex lesions. *Brain Res.* In press.

Green M. and Blake R. (1981) Phase effects in monoptic and dichoptic temporal integration: flicker and motion detection. *Vison Res.* **21**, 365–372.

Green M. (1983) Contrast detection and direction discrimination of drifting gratings. *Vision Res.* **23**, 281–289.

Gregory R. L. (1966) *Eye and Brain.* McGraw-Hill, New York.

Gregory R. L. (1970) *The Intelligent Eye,* Chap. 3, *Ambiguous, Paradoxical and Uncertain Figures.* McGraw-Hill, New York.

Grusser O.-J. and Grusser-Cornehls U. (1973) Neuronal mechanisms of visual movement perception and some psychophysical and behavioral correlations. In *Handbook of Sensory Physiology,* Vol. VII 3, *Central Visual Information B* (Edited by Jung R.), p. 333.

Hadani I., Ishai G. and Gur M. (1980) Visual stability and space perception in monocular vision: mathematical model. *J. opt. Soc. Am.* **70**, 60–65.

Hammond P. (1981) Simultaneous determination of directional tuning of complex cells in cat striate cortex for bar and texture motion. *Expl Brain Res.* **41**, 364–369.

Hammond P. and Smith A. T. (1982) On the sensitivity of complex cells in feline striate cortex to relative motion. *Expl Brain Res.* **47**, 457–460.

Helmholtz H. V. (1910) *Treatise on Physiological Optics* (Translated by Southall J. P., 1925). Dover, New York.

Henn V., Cohen B. and Young L. R. (1980) Visualvestibular interaction in motion perception and the generation of nystagmus. *Neurosci. Res. Prog. Bull.* **18**, 459–651.

Hicks T. P., Lee B. B. and Vidyasagar T. R. (1983) The responses of cells in macaque lateral geniculate nucleus to sinusoidal gratings. *J. Physiol.* **337**, 183–200.

Hildreth E. C. (1983) The measurement of visual motion. Thesis, MIT, Cambridge, MA.

Hochstein S. and Shapley R. M. (1976) Linear and non-linear spatial subunits in Y cat retinal ganglion cells. *J. Physiol.* **262**, 265–284.

Hoffman K. P. (1982) Control of the optokinetic reflex by the nucleus of the optic tract in the cat. In *Spatially Oriented Behavior* (Edited by Hein A. and Jeannerod M.), pp. 135–154. Springer, Berlin.

Hoffman K. P. (1982) Cortical versus subcortical contributions to the optokinetic reflex in the cat. In *Functional Basis of Ocular Motility Disorder* (Edited by Lennerstrand G.), pp. 303–310. Pergamon Press, Oxford.

Holub R. A. and Morton-Gibson M. (1981) Response of visual cortical neurons of the cat to moving sinusoidal gratings: response-contrast functions and spatiotemporal interactions. *J. Neurophysiol.* **46**, 1244–1259.

Horn B. K. P. and Schunk B. G. (1981) Determining optical flow. *Artificial Intelligence* **17**, 185–203.

Hoyle F. (1957) *The Black Cloud.* Harper & Row, New York.

Hubel D. H. and Wiesel T. N. (1965) Receptive fields and functional architecture in two non-striate visual areas (18 and 19) of the cat. *J. Neurophysiol.* **28**, 229–289.

Hubel D. H. and Wiesel T. N. (1962) Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *J. Physiol., Lond.* **160**, 106–154.

Johansson G. (1971) Studies on visual perception of locomotion. *Perception* **6**, 365–376.

Julesz B. (1960) Binocular depth perception of computer generated patterns. *Bell Syst. tech. J.* **39**, 1125–1162.

Julesz B. (1971) *Foundations of Cyclopean Perception.* Univ. of Chicago Press, Chicago, IL.

Julesz B. (1981) Textons, the elements of texture perception, and their interactions. *Nature* **290**, 91–97.

Julesz B. and Hesse R. I. (1970) Inability to perceive the direction of rotation of line segments. *Nature* **225**, 243–244.

Kaplan E. and Shapley R. M. (1982) X and Y cells in the lateral geniculate nucleus of macaque monkeys. *J. Physiol.* **330**, 125–143.

Keesey U. (1960) Effects of involuntary eye movements on visual acuity. *J. opt. soc. Am.* **50**, 769–774.

Keck M., Pallella T. D. and Pantle A. (1976) Motion aftereffect as a function of the contrast of sinusoidal gratings. *Vision Res.* **16**, 187.

Kelly D. H. (1979) Motion and vision. II. Stabilized spatio-temporal threshold surface. *J. opt. Soc. Am.* **69**, 1340–1349.

Kelly D. H. (1982) Fourier components of moving gratings. *Behav. Res. Meth. Instrum.* **14**, 435–437.

Kiang N. Y.-S. (1965) Discharge of single fibers in the cats

auditory nerve. *Research Monograph 35. MIT Press,* Cambridge, MA.

Kinchla R. A. and Allan L. G. (1969) A theory of visual movement perception. *Psychol. Rev.* **76,** 537–558.

Koenderink J. J. (1983) Personal communication.

Koenderink J. J. and Van Doorn A. J. (1976) Local structure of movement parallax of the plane. *J. opt. Soc. Am.* **66,** 717–723.

Korte A. (1915) Kinematoskopische Untersuchungen. *Z. Psychol.* **72,** 193–296.

Kowler E. and McKee S. P. (1984) The precision of smooth pursuit. *Invest. Ophthal. visual Sci.* **25,** 262.

Kronauer R., Daugman J. G., Zeevi Y. Y. and Namioka M. E. (1983) Fourier synthesis of spatio-temporal textures: perceptual complexity, coherence, and chaos, *Invest. Ophthal. visual Sci.* **24,** 237.

Kuffler S. W. (1952) Neurons in the retina: organization, inhibition and excitation problems. *Cold Spring Harb. Symp. quant. Biol.* **17,** 281–292.

Lange H. de (1952) Experiments on flicker and some calculations on an electrical analogue of the fovea systems. *Physica* **18,** 935–950.

Lappin J. S. and Bell H. H. (1972) Perceptual differentiation of sequential visual patterns. *Percept Psychophys.* **12,** 129–134.

Lappin J. S. and Bell H. H. (1976) The detection of coherence in moving random dots patterns. *Vision Res.* **16,** 161–168.

Lappin J. S. and Fuqua M. (1982) Non-linear recruitment in the visual detection of moving patterns. *Invest. Ophthal. visual Sci., Suppl.* **22,** 123.

Lee D. N. (1976) A theory of visual control of braking based on information about time-to-collision. *Perception* **5,** 437–457.

Lee D. N. and Aronson E. (1974) Visual proprioceptive control of standing in human infants. *Percept. Psychophys.* **15,** 529–532.

Lee D. and Lishman R. (1975) Visual proprioceptive control of stance. *J. Hum. Movmt Studies* **1,** 87–95.

Legge G. and Campbell F. W. (1981) Displacement detection in human vision. *Vision Res.* **21,** 205–214.

Lelkins A. M. M. and Koenderink J. J. (1984) Illusory motion in visual displays. *J. opt. Soc. Am.*

Lennie P. (1980) Parallel visual pathways: a review. *Vision Res.* **20,** 561–595.

Lennie P., Derrington A. M. and Krauskopf J. (1984) Spatial and chromatic mechanisms in macaque's lateral geniculate nucleus. *Invest. Ophthal. visual Sci.* **25,** 164.

Lettvin J. Y., Maturana H., McCulloch W. S. and Pitts W. H. (1959) What the frog's eye tells the frog's brain. *Proc. I.R.E.* **47,** 1940–1959.

Levinson E. and Sekuler R. (1975) The independence of channels in human vision selective for direction of movement. *J. Physiol.* **250,** 347–366.

Levinson E. and Sekuler R. (1976) Adaptation alters perceived direction of motion. *Vision Res.* **16,** 779–781.

Lisberger S. G., Evinger C., Johanson G. W. and Fuchs A. F. (1981) Relationship between eye acceleration and retinal image velocity during foveal smooth pursuit in man and monkey. *J. Neurophysiol.* **46,** 229–249.

Lisberger S. G. and Westbrook L. E. Properties of visual stimuli for the initiation of horizontal smooth pursuit eye movements. To be published.

Lishman J. R. and Lee D. N. (1973) The autonomy of visual kinaesthesis. *Perception* **2,** 287–294.

Livingston M. S. and Hubel D. H. (1984) Anatomy and physiology of a color system in the primate visual cortex. *J. Neurosci.* **4,** 309–356.

Longuet-Higgins H. C. and Prazdny K. (1980) The interpretation of moving retinal images. *Proc. R. Soc. Lond. B* **208,** 385–387.

Loomis J. and Nakayama K. (1973) A velocity analogue of brightness contrast. *Perception* **2,** 425–428.

Mansfield R. J. W. and Nachmias J. (1981) Perceived direction of motion under retinal image stabilization. *Vision Res.* **21,** 1423–1425.

Marcelja S. (1980) Mathematical description of the responses of simple cortical cells. *J. opt. Soc. Am.* **70,** 1297–1300.

Marr D. and Hildreth E. C. (1980) Theory of edge detection. *Proc. R. Soc. Lond.* **200,** 269–294.

Marr D. and Ullman S. (1981) Directional selectivity and its use in early visual processing. *Proc. R. Soc. Lond. B* **211,** 151–180.

Marr D. (1982) *Vision.* Freeman, San Francisco, CA.

Maunsell J. H. R. and Van Essen D. C. (1983a) Response properties of single units in middle temporal visual area of the macaque. *J. Neurophysiol.* **49,** 1127–1147.

Maunsell J. H. R. and Van Essen D. C. (1983b) The connection of the middle temporal visual area (MT) and their relationship to a cortical hierarchy in the macaque monkey. *J. Neurosci.* **3,** 2563–2586.

McCann J. J., Savoy R. L., Hall J. A. and Scarpetti J. J. (1974) Visibility of continuous luminance gradients. *Vision Res.* **14,** 917–927.

McKee S. P. (1981) A local mechanism for differential velocity detection. *Vision Res.* **21,** 491–500.

McKee S. P. and Nakayama K. (1984) The detection of motion in the peripheral visual field. *Vision Res.* **24,** 25–32.

McKee S. P., Silverman G. H. and Nakayama K. (1984) Velocity discrimination is not affected by random variations in temporal frequency. *Invest. Ophthal. visual Sci.* **25,** 13.

McKee S. P. and Welch L. (1985) Sequential recruitment in the discrimination of velocity. *J. opt. Soc. Am. A* **2,** 243–251.

Michael C. R. (1972) Functional organization of cells in superior colliculus of the ground squirrel. *J. Physiol.* **35,** 833.

Morgan M. J. (1980) Analogue models of motion perception. *Phil. Trans. R. Soc. Lond. B* **290,** 117–135.

Morgan M. J. and Ward R. (1980) Conditions for motion flow in dynamic visual noise. *Vision Res.* **20,** 431–435.

Motter B. C. and Mountcastle V. B. (1981) The functional properties of the light-sensitive neurons of the posterior parietal cortex studied in waking monkeys: foveal spacing and opponent vector organization. *J. Neurosci.* **1,** 3–26.

Movshon J. A., Thompson I. D. and Tolhurst D. J. (1978) Receptive field organization of complex cells in the cat's striate cortex. *J. Physiol.* **283,** 79–99.

Movshon J. A., Davis E. T. and Adelson E. H. (1980) Directional movement selectivity in cortical complex cells. *Neurosci. Abstr.* 230.

Movshon J. A., Adelson E. H., Gizzi M. and Newsome W. T. (1984) The analysis of moving visual patterns. *Expl Brain Res.*

Naegele J. R. and Held R. (1982) The postnatal development of monocular optokinetic nystagmus in infants. *Vision Res.* **22,** 341–346.

Nakayama K. (1981) Differential motion hyperacuity under conditions of common image motion. *Vision Res.* **21,** 1475–1482.

Nakayama K. (1983) Motion parallax sensitivity and space perception. In *Spatial Oriented Behavior* (Edited by Hein A. and Jeannerod M.), pp. 223–242. Springer, Berlin.

Nakayama K. (1985) Higher order derivatives of the optical velocity vector field: limitations imposed by biological hardware. In *Brain Mechanisms and Spatial Vision* (Edited by Ingle D., Jeannerod M. and Lee D.). Martinus Nijoff, Holland. In press.

Nakayama K. and Loomis J. M. (1974) Optical velocity

patterns, velocity-sensitive neurons, and space perception: a hypothesis. *Perception* 3, 63–80.

Nakayama K. and Silverman G. H. (1983) Perception of moving sinusoidal lines. *J. opt. Soc. Am.* 72.

Nakayama K. and Silverman G. H. (1984) Temporal and spatial characteristics of the upper displacement limit for motion in random dots. *Vision Res.* 24, 293–299.

Nakayama K. and Silverman G. H. (1985) Detection and discrimination of sinusoidal grating displacements. *J. opt. Soc. Am. A* 2, 267–274.

Nakayama K. and Tyler C. W. (1978) Relative motion induced between stationary lines. *Vision Res.* 18, 1663–1668.

Nakayama K. and Tyler C. W. (1981) Psychophysical isolation of movement sensitivity by removal of familiar position cues. *Vision Res.* 21, 427–433.

Nakayama K., Silverman G., MacLeod D. I. A. and Mulligan J. (1984) Sensitivity to shearing and compressive motion in random dots. *Perception*.

Narayan V., Lakshminrayan V. and Schor C. (1982) Short but not long range apparent motion stimulates OKN, OKAN, and vection. *Invest. Ophthal. visual Sci., Suppl.* 22, 124.

Newsome W. T., Gizzi M. S. and Movshon J. A. (1983) Spatial and temporal properties of neurons in macaque MT. *Invest. Ophthal. visual Sci.* 24, 106.

Newsome W. T., Wurtz R. H., Durstelor M. R. and Mikami (1983) Deficits in pursuit eye movements after chemical lesions of motion-related visual areas in the superior temporal sulcus of the macaque monkey. *Neurosci. Abstr.* 9, 154.

Orban G. A., Kennedy H. and Maes H. (1981) Response to movement of neurons in areas 17 and 18 of the cat: velocity sensitivity. *J. Neurophysiol.* 45, 1043–1058.

Orban G. A., DeWolf J. and Maes H. (1984) Factors influencing velocity encoding in the human visual system. *Vision Res.* 24, 33–39.

Pantle A. (1974) Motion after-effect magnitude as a measure of the spatio-temporal response properties of direction sensitive analyzers. *Vision Res.* 14, 1229–1236.

Pantle A. (1978) Temporal frequency response characteristics of motion channels measured with three different psychophysical techniques. *Percept. Psychophys* 24, 285–294.

Pantle A., Lehmkuhle S. and Caudill M. (1978) On the capacity of directionally selective mechanisms to encode different dimensions of moving stimuli. *Perception* 7, 261–267.

Pasternak T. and Merigan W. H. (1980) Movement detection by cats: invariance with direction and target configuration. *J. comp. Physiol. Psychol.* 94, 943–952.

Pasternak T., Shumer R., Gizzi M. and Movshon J. A. (1984) Abolition of cortical directional selectivity impairs directional discrimination in cats. *Neurosci. Abstr.* 10.

Petersik J. T., Pofahl R. and Krasnoff E. (1983) Failure to find an absolute retinal limit of a putative short range process in apparent motion. *Vision Res.* 23, 1663–1670.

Pettigrew J. D. (1973) Binocular neurons which signal change of disparity in area 18 of the cat visual cortex. *Nature* 241, 124–125.

Pola J. and Wyatt H. (1980) Target position and velocity: the stimuli for pursuit eye movements. *Vision Res.* 20, 523–534.

Pollack I. (1972) Detection of changes in spatial position: short term visual memory or motion perception? *Percept. Psychophys.* 11, 17–27.

Pollen D. A. and Ronner S. F. (1975) Periodic excitability changes across the receptive fields of complex cells in the striate and parastriate cortex of the cat. *J. Physiol.* 245, 667–697.

Pollen D. A. and Ronner S. F. (1981) Phase relationship between adjacent simple cells in the visual cortex. *Science* 212, 1409–1411.

Post R. B. and Leibowitz H. W. (1982) The effect of convergence on the vestibulo-ocular reflex and implications for perceived movement. *Vision Res.* 22, 461–465.

Prazdny K. (1981) Determining the instantaneous direction of motion from optical flow generated by a curvilinearly moving observer. *Comput. Graphics Image Process* 17, 238–248.

Purkinje J. (1825) *Beobachtungen and Versuche zur Physiologie der Sinne*, Bd II, 60.

Ramachandran V. S. and Gregory R. L. (1978) Does color provide an input to human motion perception? *Nature* 275, 55–56.

Rashbass C. (1961) The relationship between saccadic and smooth tracking eye movements. *J. Physiol.* 159, 326–338.

Regan D. (1982) Visual information channeling in normal and disordered vision. *Psychol. Rev.* 89, 407–444.

Regan D. and Beverley K. I. (1978) Looming detectors in the human visual pathway. *Vision Res.* 18, 415–421.

Regan D. and Beverley K. I. (1983) Visual fields for frontal plane motion and changing size. *Vision Res.* 23, 673–676.

Regan D. and Beverley K. I. (1984) Figure-ground segregation by motion contrast and by luminance contrast. *J. opt. Soc. Am.* 1, 433–442.

Reichardt W. (1961) Autocorrelation, a principle for the evaluation of sensory information by the central nervous system. In *Sensory Communication* (Edited by Rosenblith W. A.). Wiley, New York.

Reichardt W. and Poggio T. (1979) Figure-ground discrimination by relative movement in the visual system of the fly. *Biol. Cybernet.* 35, 81–100.

Richards W. (1971) Motion detection in man and other animals. *Brain Behav. Evol.* 4, 162–181.

Richards W. and Regan D. (1973) A stereo field map with implications for disparity processing. *Invest. Ophthal. visual Sci.* 12, 904–909.

Richards W., Rubin J. and Hoffman D. O. (1982) Equation counting and the interpretation of sensory data. *Perception* 11, 557–576.

Riggs L. A., Ratliff F., Cornsweet J. C. and Cornsweet T. N. (1953) The disappearance of steadily fixated visual test objects. *J. opt. Soc. Am.* 43, 495–501.

Robinson D. A. (1965) The mechanics of human smooth pursuit eye movement. *J. Physiol.* 180, 569–591.

Robinson D. A. (1968) Eye movement control in primates. *Science* 161, 1219–1224.

Robinson D. A. (1982) The use of matrices in analyzing the three dimensional behavior of the vestibulo-ocular reflex. *Biol. Cybernet.* 46, 53–66.

Rogers B. and Graham M. (1979) Motion parallax as an independent cue for depth perception. 8, 125–134.

Rogers B. J. and Graham M. (1982) Similarities between motion parallax and stereopsis in human depth perception. *Vision Res.* 22, 261–270.

Rogowitz B. E. (1984) The breakdown of size constancy under stroboscopic illumination. In *Sensory Experience, Adaptation and Perception, Festschrift for Ivo Kohler* (Edited by Spillman L. and Wooten B. R.), pp. 201–213. Lawrence Erlbaum, Hillsdale, NJ.

Sakitt B. and Barlow H. B. (1982) A model for the economical encoding of visual image in cerebral cortex. *Biol. Cybernet.* 43, 97–108.

van Santen J. P. H. and Sperling G. (1984) A temporal covariance model of human motion perception. *J. opt. Soc. Am. A* 1, 451–473.

Schift W. and Detwiler M. L. (1979) Information used in judging impending collision. *Perception* 8, 647–658.

Schiller P. H., Finlay B. L. and Volman S. F. (1976) Quantitative studies of single-cell properties in monkey striate cortex. I. Spatiotemporal organization of receptive fields. *J. Neurophysiol.* 39, 1288–1319.

Schiller P. H. and Malpeli J. G. (1978) Functional specificity of lateral geniculate nucleus laminae of the

rhesus monkey. *J. Neurophysiol.* **41**, 788-797.

Schor C. and Narayan V. (1981) The influence of field size upon the spatial frequency response of optokinetic nystagmus. *Vision Res.* **21**, 985-994.

Scobey R. P. and Horowitz J. M. (1976) Detection of image displacement by phasic cells in peripheral visual fields of the monkey. *Vision Res.* **16**, 15-24.

Sekuler R. W. and Ganz L. (1963) After-effects of seen motion with a stabilized retinal image. *Science* **139**, 419-420.

Sekuler R. (1975) Visual motion perception. In *Handbook of Perception* (Edited by Carteretta E. and Friedman M.) Vol. 5, Chap. 11.

Sekuler R., Pantle A. and Levinson E. (1978) Physiological basis of motion perception. In *Handbook of Sensory Physiology* (Edited by Held R., Leibowitz H. and Teuber H.), Vol. III, pp. 67-96.

Shaw E. (1978) Schooling fishes. *Am. Scient.* **66**, 166-175.

Sillito A. M. (1977) Inhibitory processes underlying the directional specificity of simple, complex and hypercomplex cells in the cat's visual cortex. *J. Physiol.* **271**, 699-720.

Simpson J. I., Graf W. and Leonard C. (1981) The coordinate system of visual climbing fibers to the flocculus. In *Progress in Oculomotor Research* (Edited by Fuchs A. F. and Becker W.), pp. 475-484. Elsevier, Amsterdam.

Spear P. D. and Baumann T. P. (1975) Receptive field characteristics of single neurons in lateral suprasylvian visual area of the cat. *J. Neurophysiol.* **38**, 1403-1420.

Sperling G. (1976) Movement perception in computer-driven visual displays. *Behav. Res. Instrum.* **8**, 144-151.

Srinivasan M. V. (1977) A visually evoked roll response in the housefly. *J. comp. Physiol.* **119**, 1-14.

Srinivasan M. V. (1983) The impulse response function of a movement detecting neuron and its interpretation. *Vision Res.* **23**, 659-664.

Srinivasan M. V. and Dvorak D. R. (1979) The waterfall illusion in an insect visual system. *Vision Res.* **19**, 1435-1437.

Sterling P. and Wickelgren B. G. (1969) Visual receptive fields in the superior colliculus of the cat. *J. Neurophysiol.* **32**, 1-15.

Suzuki D. A., Noda H. and Kase M. (1981) Visual and pursuit eye movement related activity in posterior vermis of monkey cerebellum. *J. Neurophysiol.* **46**, 1120-1139.

Suzuki D. A. and Keller E. L. (1984) Visual signals in the dorsolateral pontine nucleus of the alert monkey: their relationship to smooth-pursuit eye movements. *Expl Brain Res.* **53**, 473-478.

Thompson P. (1984) The encoding of velocity of movement in the human visual system. *Vision Res.* **24**, 41-46.

Thorson J. (1966a) Small-signal analysis of a visual reflex in the locust. I. Input parameters. *Kybernetik* **3**, 41-53.

Thorson J. (1966b) Small-signal analysis of a visual reflex in the locust. II. Frequency dependence. *Kybernetik* **3**, 53-66.

Todd J. T. (1981) Visual information about moving objects. *J. exp. Psychol. Hum. Percept. Perform.* **7**, 795-810.

Tolhurst D. J. (1973) Separate channels for the analysis of shape and the movement of a moving visual stimulus. *J. Physiol.* **231**, 385-402.

Turvey M. and Remez R. E. (1979) Visual control of locomotion in animals: an overview. *Proc. Conf. Inter-relations of Communicative Sciences* (Edited by Harmon L.), National Science Foundation Publication, Washington, DC.

Tyler C. W. (1973) Periodic vernier acuity. *J. Physiol.* **228**, 637-647.

Tyler C. W. (1974) Induced stereomovement. *Vision Res.* **14**, 609-613.

Tyler C. W. and Torres J. (1972) Frequency response

characteristics for sinusoidal movement in the fovea and periphery. *Percept. Psychophys.* **12**, 232-236.

Ullman S. (1979) The interpretation of structure from motion. *Proc. R. Soc. Lond.* B **203**, 405-426.

Ullman S. (1983) The measurement of visual motion. *Trends in Neurosciences.* Elsevier, Amsterdam.

Van der Horst G. J. C. and Bouman M. A. (1969) Spatio-temporal chromaticity discrimination. *J. opt. Soc. Am.* **59**, 1482-1488.

Van Doorn A. J. and Koenderink J. J. (1982a) Spatial properties of the visual detectability of moving spatial white noise. *Expl Brain Res.* **45**, 189-195.

Van Doorn A. J. and Koenderink J. J. (1982b) Temporal properties of the visual detectability of moving spatial white noise. *Expl Brain Res.* **45**, 179-182.

Van Doorn A. J. and Koenderink J. J. (1982c) Visibility of moving gradients. *Biol. Cybernet.* **44**, 167-175.

Van Doorn A. J. and Koenderink J. J. (1983) Detectability of velocity gradients in moving random dot patterns. *Vision Res.* **23**, 799-804.

Van Doorn A. J. and Koenderink J. J. (1984) The structure of the human motion detecting system. *Proc. I.E.E.E.*

Van Essen D. C. (1979) Visual areas of the mammalian cerebral cortex. *A. Rev. Neurosci.* **2**, 227-263.

Van Nes F. L., Koenderink J. J., Nas H. and Bouman M. (1967) Spatio-temporal modulation transfer in the human eye. *J. opt. Soc. Am.* **57**, 1082-1088.

Virsik R. and Reichardt W. (1976) Detection and tracking of moving objects by the fly *Musca domestica*. *Biol. Cybernet.* **23**, 83-98.

Wallach H. (1935) Über visuell wahrgenommene Bewegungsrichtung. *Psychol. Forschung.* **20**, 325-380.

Wallach H. and O'Connell D. N. (1953) The kinetic depth effect. *J. exp. Psychol.* **45**, 205-217.

Watanabe A., Mori T., Nagata S. and Hiwatashi K. (1968) Spatial sine wave responses of the human visual system. *Vision Res.* **8**, 1245-1263.

Watson A. B. and Ahumada A. J. (1983) A look at motion in the frequency domain. *NASA Tech. Memorandum* 84352.

Watson A. B. and Ahumada A. J. (1984) A model of how humans sense image motion. *Invest. Ophthal. visual Sci.* **25**, 14.

Watson A. B., Thompson P. G., Murphy B. J. and Nachmias J. (1980) Summation and discrimination of gratings moving in opposite directions. *Vision Res.* **20**, 341-347.

Watson A. B., Ahumada A., Farrell J. E. (1983) The window of visibility: a psychophysical theory of fidelity in time-sampled visual motion displays. NASA Tech. Paper 2211.

Wertheimer G. (1912) Experimentelle Studien über das Sehen von Bewegungen. *Z. Physchol.* **61**, 161-265.

Westheimer G. (1978) Spatial phase sensitivity for sinusoidal grating targets. *Vision Res.* **18**, 1073-1074.

Westheimer G. (1979) The spatial sense of the eye. *Invest. Ophthal. visual Sci.* **18**, 893-912.

Westheimer G. and McKee S. P. (1975) Visual acuity in the presence of retinal image motion. *J. opt. Soc. Am.* **65**, 847-850.

Wilson H. R. and Bergen J. R. (1979) A four mechanism model for spatial vision. *Vision Res.* **19**, 19-32.

Wohlgemuth A. (1911) On the aftereffect of seen movement. *Br. J. Psychol. Monogr., Suppl.*

Wyatt H. and Daw N. W. (1976) Specific effects of neurotransmitter antagonists on ganglion cells in rabbit retina. *Science* **204**, 205.

Yasui S. and Young L. R. (1975) Perceived visual motion as effective stimulus to pursuit eye movement system. *Science* **190**, 906-908.

Zeki S. M. (1974a) Functional organization of a visual area in the posterior bank of the superior temporal sulcus of the rhesus monkey. *J. Physiol., Lond.* **236**, 549-573.

Zeki S. M. (1974b) Cells responding to changing image size and disparity in the cortex of the rhesus monkey. *J. Physiol., Lond.* **242**, 827–841.

Zeki S. (1980) The response properties of cells in the middle temporal area (area MT) of owl monkey visual cortex. *Proc. R. Soc. Lond. B* **207**, 239–248.

Zihl J., Cramon D. von and Mai N. (1983) Selective disturbance of movement vision after bilateral brain damage. *Brain* **106**, 313–340.