Pergamon

0042-6989(95)00121-2

Vision Res. Vol. 36, No. 3, pp. 471–490, 1996

0042-6989/96 $15.00 + 0.00

Ideal Observer for Heading Judgments

JAMES A. CROWELL,* MARTIN S. BANKS*†

Received 18 July 1994; in revised form 13 February 1995; in final form 10 April 1995

Several aspects of the viewing situation affect the ability to determine heading from optical flow. These include the amount of depth variation and number of texture elements in the scene, the location and amount of the visual field stimulated, and the position of the focus of expansion within the stimulus. Without a quantification of the discrimination information provided by the stimuli presented to the observer, it is impossible to determine how much of an observed change in performance reflects the properties of neural mechanisms and strategies employed by the observer. To enable a better quantification, we developed an ideal observer for the discrimination of heading from random-dot flow fields. Internal noises of the ideal observer were set by the results of single-dot velocity discrimination experiments. We compared human and ideal observer performance in discriminating headings with different patterns of flow (e.g. radial vs laminar) presented on different parts of the retina. Efficiency—the ratio of ideal and human thresholds—was fairly constant for the various flow patterns and retinal eccentricities. This outcome indicates that most of the variation in human observers' ability to estimate heading from the flow patterns and retinal loci considered here is due to changes in the discrimination information provided by the stimulus after measurement by the visual system. In the discussion, we show how the ideal observer can be used to quantify the spatial distribution of heading discrimination information for any observer translation through any scene represented by dots.

Optic flow Motion Ideal observer Signal detection Heading

INTRODUCTION

As an observer moves through the environment, a changing pattern of light falls on the retina. Gibson (1950, 1966) called this pattern the optic flow field and showed that it could provide information to guide navigation by allowing the observer to estimate the direction of self-motion with respect to environmental landmarks (Fig. 1). There is now substantial psychophysical evidence that human observers can use optic flow to judge their direction of self-motion, or heading, in a variety of situations (Warren & Hannon, 1988, 1990; Royden, Banks & Crowell, 1992). However, as in most visual tasks, the accuracy of performance depends on several stimulus properties; these include the number of elements presented, the field of view, the retinal region of stimulation, and the type of flow (Rieger & Toet, 1985; Warren, Morris & Kalish, 1988; Warren & Kurtz, 1992; Crowell & Banks, 1993a, b).

Psychophysical experiments measure the performance of the visual system as a whole, so the observation that a change in stimulus properties leads to a change in performance does not yield a simple interpretation. In the case of heading judgments, two possibilities come to mind: (1) the psychophysical effect could reflect a change in the amount of information available in the internal representation of the optic flow field; or (2) it could reflect a change in the efficiency with which the heading is computed from that internal representation. Here we present a means to distinguish between these two interpretations.

The importance of making such a distinction can hardly be overstated: In the study of 3D object- or self-motion, psychophysical results will hopefully reveal the properties of the mechanisms, cues, or strategies human observers use to perform various tasks. We cannot conclude, however, that an observed psychophysical effect is due to such properties unless we can rule out the possibility that the effect is simply a consequence of changing the amount of information provided by the experimental displays.

The following example will help to make this point clear. Numerous investigators have examined the ability to maintain or estimate direction of self-motion in simulated flight. For example, there have been many studies of the ability to discriminate changes in altitude or to maintain a fixed altitude in simulated flight above a ground plane (Warren, 1982). Observers are generally better at maintaining altitude or detecting altitude change when the texture in the display consists of lines parallel to the component of forward motion (i.e. the component parallel to the ground) than when the texture consists of lines perpendicular to the forward component (Flach, Hagen & Larish, 1992; Warren, 1988; Wolpert, 1987, 1988). It is not obvious why performance should

*School of Optometry and Department of Psychology, University of California, Berkeley CA 94720-2020, U.S.A. [*Email* marty@john.berkeley.edu].

†To whom all correspondence should be addressed.

be better in the first case because altitude changes in both cases cause changes in the flow field. If the lines are parallel to the forward motion component, a decrease in altitude causes their images to rotate towards the horizontal. If the lines are perpendicular to the forward component, an altitude decrease causes changes in the speed of the images of the lines. There are two possible interpretations for the superiority of lines parallel to the observer's motion. First, human observers may have developed more efficient neural mechanisms or strategies for picking up altitude information from the projections of such lines over time (such as optical splay angle, the orientation of the lines' images with respect to the vertical meridian). Second, changes in lines parallel to the motion may simply provide better information for the specification of an altitude change without reference to the type of nervous system to which the information is presented; that is, any efficient machine might exhibit the same behavior. For example, when the lines are perpendicular to the forward motion, the observer must detect a change on top of the optic flow created by the forward motion (with parallel lines, forward motion creates no optic flow) and this might lead to poorer performance. Without a means to decide which is the better interpretation, one does not know whether these results tell us more about the observer or about the information in the displays.

Another example allows us to introduce the experimental observations upon which our analysis is based. While walking or operating a motor vehicle, people usually fixate in a direction near their path of motion, but it is not uncommon to look in another direction. When a person looks near the path (holding gaze position fixed relative to the direction of self-motion by looking at a distant object), the central visual field is exposed to a radial pattern of flow with the focus of expansion near the fovea. When a person looks in another direction (again holding gaze position relative to self-motion fixed), the central visual field is presented with a more laminar flow pattern and the focus of expansion is imaged on the peripheral retina. Does the ability to estimate one's heading differ in those two situations? Crowell and Banks (1993a), Warren and Kurtz (1992), and Wolpert (1987) examined this question psychophysically by presenting different types of flow fields to different parts of the retina.

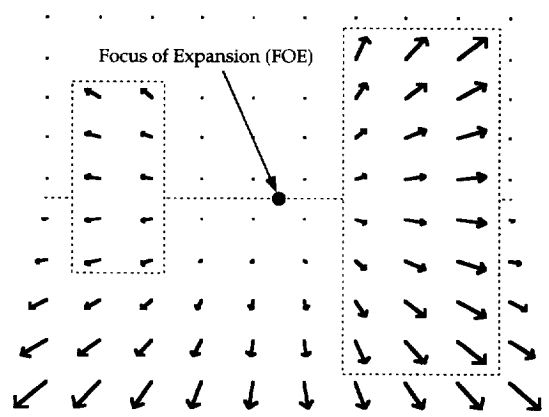


FIGURE 1. Flow field created by observer translation. The scene is a ground plane, a near wall on the right and a farther wall on the left. The vectors represent the motions in the image plane of dots in the scene. The observer's heading is indicated by the circle, and it can be estimated by triangulation using two or more non-collinear flow vectors.

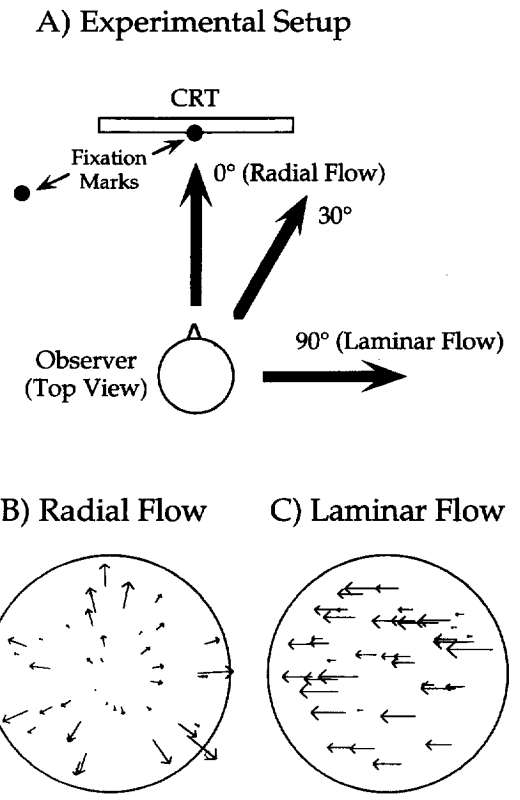


FIGURE 2. Schematic of the experiments of Crowell and Banks (1993a). (A) Top view of the experimental apparatus. Retinal eccentricity was varied by moving the fixation point (represented by ●); the type of flow presented was manipulated by varying the observer's heading with respect to the display, or *heading eccentricity*. The simulated heading is represented by the arrows. (B) Radial flow field created by translation towards the display (heading eccentricity = 0 deg). (C) Laminar flow field created by translation parallel to the display (heading eccentricity = 90 deg).

Crowell and Banks (1993a) asked observers to discriminate directions of self-motion when presented with small flow fields. When the focus of expansion was in the middle of the stimulus, a radial flow field was created [Fig. 2(B)]; when the focus was not within the stimulus, a laminar field was created [Fig. 2(C)]. Crowell and Banks reported that heading discrimination thresholds are much lower with radial than with laminar flow fields. Should this difference in performance be attributed to a greater amount of discrimination information provided by radial flow or to more efficient use of the information specifying heading with radial flow? The approach presented here allows one to distinguish these possibilities.

The approach, based on the theory of ideal observers (Green & Swets, 1966), allows one to quantify the performance limitations imposed by the discrimination information in the stimulus. As applied to vision by Barlow (1958, 1962), Geisler (1989), Rose (1948), and others, the approach consists of comparisons of human

observers' performance with ideal observer performance for the same stimuli. Ideal observers have three main elements: (1) a precise description of the stimuli including any random variation or noise; (2) descriptions of how the stimuli are modified by the visual processing stages incorporated in the analysis; and (3) a statistically optimal rule for selecting a response when presented with a sample from the stimulus distributions (the "decision rule"). Such observers specify the best possible performance (e.g. the dimmest detectable light) given the variability in the stimulus and information losses in the incorporated processing stages. Because ideal observer performance is the best possible, the thresholds obtained are a measure of the discrimination information available in the stimulus once processed by the incorporated stages. Poorer performance by the human observer must be due to information losses in processing stages that were not incorporated in the ideal observer.

An important concept in this approach is that of the efficiency of the human observer. The efficiency is a measure of the gap between human and ideal performance;* it is the number that computational models must attempt to explain.

If a given experimental manipulation leads to a change in human performance, there is a natural tendency to develop theories of the observer to explain the change. However, if the observer's efficiency is unchanged by the manipulation (because the amount of information available to make the discrimination is changing), such theories of the observer are inappropriate (Watson, 1987).

Here we present an ideal observer for heading discrimination tasks and show how comparisons of human and ideal performance allow a better understanding of the limits to performance in these tasks. Specifically, we examine whether variations in the ability to estimate heading from different types of flow fields presented on different parts of the retina should be attributed to changes in the amount of discrimination information available in the optic flow field after it has been measured by the human visual system or to changes in the efficiency with which the available information is used to estimate heading.

Heading discrimination with different retinal regions and types of flow

Before describing the ideal observer, we present the human data we wish to analyze. These data on heading discrimination for different types of flow presented at different retinal eccentricities were reported previously by Crowell and Banks (1993a).

Observers were shown sequences of dots moving so as to simulate motion of the observer in a straight line through a random 3D cloud of points. Observers viewed the stimuli monocularly through a circular aperture 10 deg in diameter. Two motion sequences were presented on each trial; the two sequences simulated different directions of self-motion, or headings [Fig. 2(A)]. The headings were always in the horizontal plane. Observers indicated whether the change in heading between the two intervals was rightward or leftward. The angular difference between the two headings was varied using a staircase procedure to find the 71% correct heading discrimination threshold. The stimuli were presented at retinal eccentricities of 0 deg (central fixation), ± 5, ± 10, and ± 40 deg; positive values represent temporal retinal stimulation and negative values nasal stimulation. The type of flow was manipulated independently by varying the *heading eccentricity*, which is the angle between the heading and the center of the stimulus. Heading eccentricities of 0 deg [radial flow; Fig. 2(B)], ± 5, ± 10, ± 40, and ± 70 deg [nearly laminar flow; [Fig. 2(C)] were presented. By independently manipulating the heading eccentricity and the retinal eccentricity, we could ask whether different parts of the retina are specialized for processing different kinds of flow fields.

Figure 3 shows the data from one observer plotted in two ways. Figure 3(A) displays heading discrimination threshold as a function of heading eccentricity for a variety of retinal eccentricities; Fig. 3(B) displays threshold as a function of retinal eccentricity for a variety of heading eccentricities. There is a large and consistent effect of heading eccentricity. The lowest threshold of about 0.2 deg (at a heading eccentricity of 0 deg and a retinal eccentricity of 0 deg) is 100 times lower than the highest threshold of about 20 deg (at a

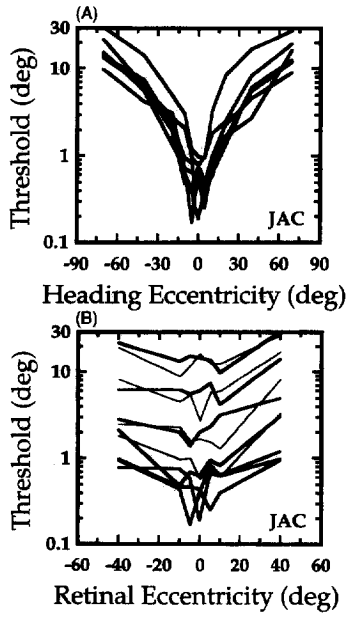


FIGURE 3. Heading thresholds from Crowell and Banks (1993a). (A) Thresholds for observer JAC are plotted as a function of heading eccentricity; heading eccentricity of 0 deg yields radial flow and 90 deg yields laminar flow. The separate functions indicate different retinal eccentricities. (B) Thresholds for observer JAC are plotted as a function of the retinal eccentricity of stimulation, a separate function for each heading eccentricity.

*Technically, the efficiency with which a human observer can discriminate two stimuli is defined as the squared discriminability (d'^2) for the human observer divided by the squared discriminability for an ideal observer. However, this quantity is often difficult to measure, particularly when it is low. We will use the term "efficiency" to refer to the ratio of the ideal discriminator model's threshold (71% correct) to the human observer's threshold.

heading eccentricity of ± 70 deg at any retinal eccentricity). The effect of retinal eccentricity is smaller and less consistent. It is mostly confined to a set of narrow dips corresponding to flow fields in which the focus of expansion falls on the observer's fovea.

Problem of interpretation. The observation that heading is perceived more precisely with radial than with laminar flow can be interpreted in two ways. First, the advantage with radial flow could reflect a neural specialization for processing such a flow pattern. Support for this hypothesis is provided by Regan and Beverley (1979, 1980) who demonstrated visual channels sensitive to looming stimuli; the existence of such channels might well yield a greater precision in perceiving heading with radial flow as compared with laminar flow. On the other hand, the effect of heading eccentricity could be a consequence of the variation in discrimination information as the distance between the focus of expansion and the center of the stimulus is varied. Support for this idea comes from Koenderink and van Doorn (1987) and Crowell and Banks (1993a); they argued on geometric grounds that random perturbations of flow vectors should have a greater effect on the precision of heading estimates when the separation between the flow vectors and the focus of expansion is large. This idea is at least qualitatively consistent with the data in Fig. 3(A).

The observation of a small advantage with radial flow in the central as opposed to the peripheral visual field can also be interpreted in two ways. First, it may be a manifestation of special mechanisms for processing radial flow centered on or near the fovea (Warren & Kurtz, 1992). Second, it may be a consequence of having receptive fields tuned to slower speeds in or near the fovea (Albright, 1984; Koenderink, van Doorn & van de Grind, 1985; McKee & Nakayama, 1984); this idea is qualitatively consistent with the data in Fig. 3(B) because radial flow patterns provide slowly-moving dots near the focus of expansion.

To interpret the data presented in Fig. 3, we compared ideal and human performance on the same heading discrimination tasks in the fashion schematized in Fig. 4.

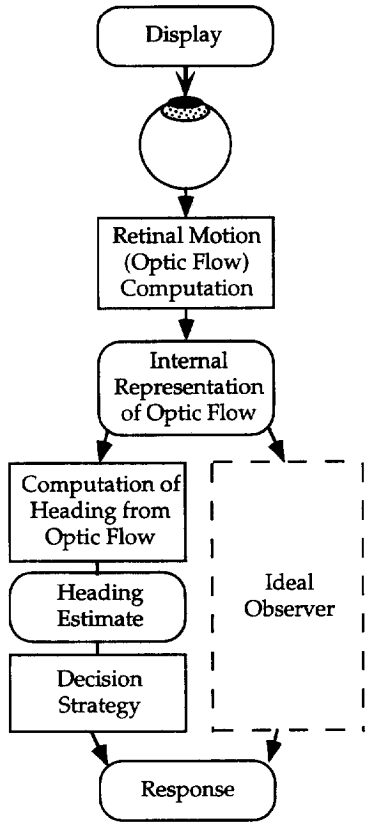


FIGURE 4. Schematic of the modified ideal observer approach. We want to compare human performance to that of an ideal observer having the same front-end including the stage at which an internal representation of motion in the retinal image (optic flow) is computed from the changing pattern of light. Any differences between human and ideal performance reflect the efficiency of the later processing stages. Lacking a quantitatively precise model of human optic flow computation, we developed the approach described in the text for quantifying the noise in the internal representation of optic flow.

Constructing the ideal observer model

An ideal observer is conceptually very simple. It consists of three elements: (1) a description of two or more nominal stimuli (signals) and of the random variation in each (noise); (2) a description of how the stimuli are affected by the processing stages incorporated in the model; and (3) a likelihood-ratio decision rule.

Constructing an ideal observer for heading discriminations is complicated by the lack of a quantitative model of how the human visual system computes motion in the retinal image (element 2 above). With such a model, we could derive an ideal observer for heading discriminations that took as its input the raw motion sequences presented in our experiments. There are many candidates (e.g. Adelson & Bergen, 1985; Fleet & Jepson, 1989; Heeger, 1987; van Santen & Sperling, 1985; Watson & Ahumada, 1985), but none is sufficiently detailed nor universally accepted. Lacking such a model, we have chosen to make a few additional assumptions about the signals and noises that are the input to the heading computation that allow us to estimate them from psychophysical data.

Representing the signals. The "signal" is the optic flow field and is represented by a vector field, each vector representing the position and velocity of a dot. The base of the vector represents the dot's initial position and the tip its angular velocity at the beginning of the motion sequence.

Incorporating noises. The noise is the random error associated with the measurement of optic flow. We made three assumptions about these measurement errors.

(1) We assumed that the errors associated with the measurement of the speed and direction of each dot are statistically independent of the measurement errors for the other dots.
(2) We assumed that the means of the noise distributions are zero, that is, that the measurements are unbiased.
(3) For the sake of simplicity and computational convenience, the speed and direction measurement

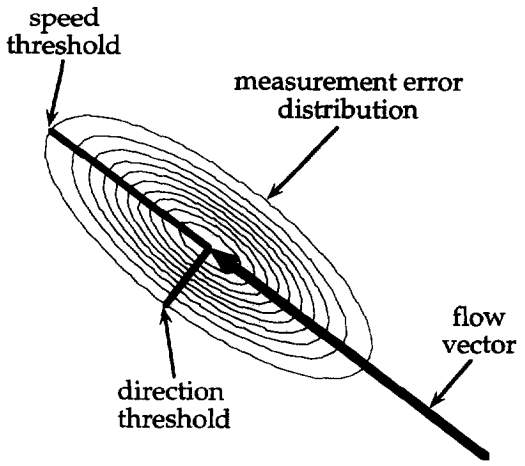


FIGURE 5. Schematic of the noise distributions associated with each flow vector. The motion of each dot was represented by a vector. The contour lines around the tip of the vector represent the flow vector measurement error probability density function; for a given motion vector, the value of this function at any point (actually, the integral of the function over a very small region) represents the probability that the measured dot motion will assume that value. The length of the distribution (σ_s) is determined from the single-dot speed discrimination thresholds and the width (σ_d) from the direction thresholds.

error distributions were assumed to be Normal.* Given means of zero, we only need to specify the standard deviation (σ) to determine each distribution completely. Standard deviations for a variety of base speeds and retinal eccentricities were estimated from the results of the single-dot direction and speed discrimination experiments described in the Appendix.

The results of the single-dot discrimination experiments are summarized in Figs 15 and 16. As expected, direction and speed discrimination thresholds were roughly constant at high speeds and increased markedly at low speeds. The speed below which threshold began to rise was higher at greater retinal eccentricities. The variations in threshold lead to corresponding variations in the standard deviations of the ideal observer's measurement error distributions (see Fig. 6).

If we represent a dot's motion by a vector, then the measurement errors can be represented by a 2D probability distribution that is (by assumption) the product of two 1D distributions: a speed error distribution parallel to the vector and an orthogonal direction error distribution (Fig. 5). The human observers' speed discrimination thresholds in the single-dot experiment are used to set the standard deviations of the speed error distributions (the lengths of the joint distributions), and their direction discrimination thresholds are used to set the standard deviations of the direction error distributions (the widths).

The speed error standard deviations were calculated using the formula $\sigma_s = \delta m_s / d'$, where δm_s is the 71% correct speed discrimination threshold (i.e., the difference between the means of the distributions that yielded 71% correct performance) and d' is the discriminability index corresponding to 71% correct ($d' = 0.78$). For most cases the direction standard deviations (σ_d) were calculated in a similar fashion from the direction discrimination data. For cases in which the threshold was high (when a wrapped Normal density function is not well-approximated by an ordinary Normal) the value of σ_d was determined by numerically integrating the

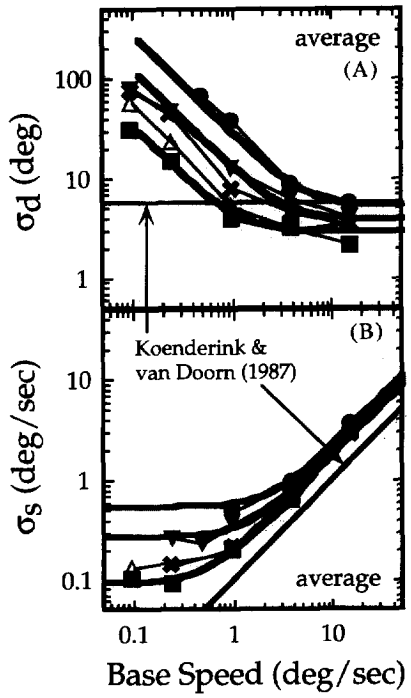


FIGURE 6. Width (σ_d) and length (σ_s) parameters of ideal observer's noise distributions as a function of base speed. The data points are the averages of the three observers' thresholds. (A) The single-dot direction discrimination data and the one-parameter curves used to fit each data set; different lines represent curves obtained at retinal eccentricities of 0, 10 and 40 deg. The curves were used to interpolate or extrapolate a σ_d-value for any combination of base speed and retinal eccentricity. (B) The single-dot speed discrimination data and the one-parameter curves used to fit each data set; again different lines represent curves obtained at different retinal eccentricities. The curves were used to interpolate or extrapolate σ_s-values for different base speeds and retinal eccentricities. In each panel the thinner black line indicates the noises assumed by Koenderink and van Doorn (1987) (see Discussion).

*For direction, the distributions were actually wrapped Normals. A wrapped Normal distribution is a linear Normal distribution that has been wrapped around the circumference of a circle with each successive wrap added on to the ones before, i.e.

$$P(\theta) = \frac{1}{\sigma\sqrt{2\pi}} \sum_{k=-\infty}^{\infty} e^{\frac{-(\theta + 2\pi k)^2}{2\sigma^2}}.$$

For small values of σ ($\sigma < \sim 70$ deg), a wrapped Normal looks like an ordinary Normal distribution; as σ goes to infinity, it asymptotically approaches a uniform distribution, $P(\theta) = 1/2\pi$ (θ in radians). A wrapped normal random variable is easy to generate; it is an ordinary Normal variable mod 2π. Wrapped Normal variables are also convenient because they possess the additive property (Mardia, 1972): that is, the sum of two random wrapped Normal variables with width parameters σ_1 and σ_2 is a wrapped Normal variable with width parameter $\sigma_{1+2} = \sqrt{\sigma_1^2 + \sigma_2^2}$. The density function is well-approximated by a Normal density for small values of σ and by the expression $P(\theta) = (1 + 2(\rho \cos(\theta) + \rho^4 \cos(2\theta)))/2\pi$ for large values of σ, where $\rho = \exp(-\sigma^2/2)$.

appropriate areas under the probability density functions for a trial value of σ_d and adjusting that value until 71% correct performance was achieved. Finally, a one-parameter family of curves was fitted to the sets of σ_s- and σ_d-values with retinal eccentricity as the parameter. Three of these curves (for retinal eccentricities of 0, 10, and 40 deg) are shown in Fig. 6 along with the σ_s- and σ_d-values calculated from the thresholds (averaged across observers) of the single-dot experiment.

The velocities of the dots in the heading experiments of Crowell and Banks (1993a) spanned a large and continuous range. Thus, the one-parameter functions shown in Fig. 6 were used to interpolate values for σ_s and σ_d in modeling those experiments. Extrapolation was occasionally necessary, but the dot speeds in the heading experiments were almost always within the range covered by the single-dot experiment.

Speed uncertainty in a 3D cloud

The sensation of self-motion is aided by presenting a stimulus with a wide range of simulated depths (Andersen & Braunstein, 1985; Howard & Heckmann, 1989). A 3D cloud of randomly-positioned dots is very effective in this regard, but it complicates the modeling. The depth (and hence the speed) of a dot in the cloud is random, and this randomness in the stimulus must be incorporated in the ideal observer. This is accomplished by treating each dot's speed (the length of each flow vector) as an uncertainty parameter. This uncertainty with regard to speed is represented by a probability density function. In the experiments of Crowell and Banks (1993a), the distances to the dots in the 3D cloud were transformed such that the distribution of velocities along any given line of sight was uniform at the beginning of a trial. Thus, the density function is zero for speeds slower than those of the most distant dots or faster than the closest ones; in between, the density function is

$$f(v_d) = \frac{1}{T\sin(\theta)\left(\frac{1}{r_f} - \frac{1}{r_b}\right)}, \left(\frac{T\sin(\theta)}{r_b} \leqslant v_d \leqslant \frac{T\sin(\theta)}{r_f}\right) \quad (1)$$

where v_d is the angular velocity of a dot in the display, T is the observer's translational speed, θ is the visual angle between the heading and the dot in question, and r_f and r_b are the radial distances from the observer to the front and back of the cloud, respectively.

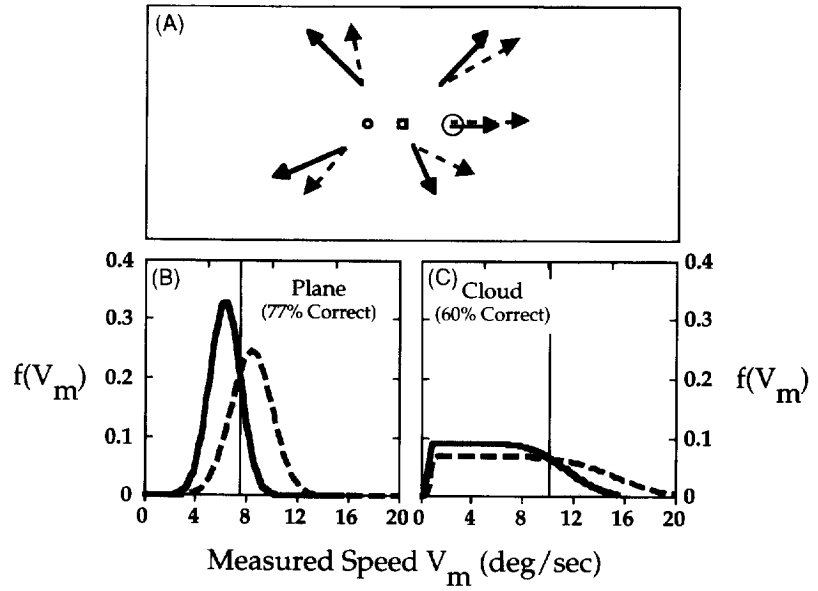
How does this affect the ideal observer calculations? The ideal observer works with the likelihoods of observing a given set of flow vectors given each of the possible headings. If all of the dots in each stimulus were at one known depth, then calculating the likelihoods would be straightforward: there would be one expected speed at each point associated with each heading, and the likelihood of any deviation from this value in the measured speed would be given directly by the value of the appropriate speed error probability density function. With a range of possible depths (and hence of possible speeds), the likelihood of any given measured speed has to be calculated by summing across the set of all possible depths at that point in the scene. Furthermore, because the direction error standard deviation (σ_d) depends on speed, a similar summation has to be performed to calculate the likelihood of any observed vector direction given each of the possible headings. In order to calculate the likelihood of a variable when there is an uncertainty parameter, the probability density function must be integrated over all possible values of that parameter. For example, the likelihood of the observed direction (α_i) for the ith flow vector given the hypothesis of the jth heading (H_j) is

$$f(\alpha_i|H_j) = \int_{\frac{T\sin(\theta)}{r_b}}^{\frac{T\sin(\theta)}{r_f}} f(\alpha_i|H_j, v_d) f(v_d) d(v_d) \quad (2)$$

where $f(\alpha_i|H_j, v_d)$ is the conditional probability of observing flow vector direction α_i given heading H_j and a speed of v_d in the display, and $f(v_d)$ is the prior distribution of display speeds [equation (1)]. We evaluated these integrals numerically using a fourth-order Runge-Kutta method (Press, Flannery, Teukolsky & Vetterling, 1988).

The existence of uncertainty about the dot speeds degrades the performance of the ideal observer model. The reason for this decrement in performance is illustrated in Fig. 7. Figure 7(A) shows flow fields created by two different headings; the corresponding foci of expansion are indicated by the small square and circle. Figure 7(B, C) shows the distributions of dot speeds that the observer can expect to see at a single location in the field, indicated by the circle in Fig. 7(A). Figure 7(B) shows the probability density function of observed speeds for a plane stimulus given each of the two possible headings. In this case, there is only a single possible speed in the stimulus for each of the two headings, corresponding to the locations of the peaks of the two functions. The two probability distributions result solely from the Normally-distributed noise added by the ideal observer's front end. The vertical line marks the criterion value of speed that optimizes the proportion of correct responses; that is, if the ideal observer used only the information in the speed of a single dot at this point in the field, it should select a response based on whether the observed speed were above or below this criterion. If it were to do so, it would achieve 77% correct performance.

Figure 7(C) shows the density functions for observed speeds with a cloud stimulus. Now the ideal observer's internal speed noise is added on top of the random variation in speed due to randomization of depths within the cloud. As a result, the distributions of observed speeds are broader and overlap more; in this situation, the best possible performance using this vector speed is only 60% correct. The same considerations apply to the other dots in the stimulus. As a consequence, the model's heading estimates would be more precise with the single plane than they are for 3D clouds. Whether such a reduction in depth (and therefore speed) uncertainty makes any difference to human observers is an important and unresolved question.



IGURE 7. Probability densities of observed speeds (after addition of internal noise) in (B) a plane stimulus and (C) a 3D cloud with the same mean speed. The two curves in each panel are for the same location in the visual field with two different headings; the location is indicated by ○ and the two headings by the ○ and □ in (A). Because there is random variability in the speed in the display in the case of a 3D cloud, the distributions for the two headings overlap more and there is less information available for discriminating the two headings; using only a single vector speed observation, the ideal observer would attain 77% correct performance with the plane, but only 60% correct with the cloud.

Implementation of optimal decision rule

For two-interval experiments like the ones considered here, an ideal discriminator operates by calculating a single number for the trial, the likelihood ratio, and then making a decision by determining whether that number is greater or less than some criterion. This decision rule yields optimal discrimination performance with respect to a variety of goals (Green & Swets, 1966). The likelihood is the probability of an observation—in our case, an observation consists of two samples of flow vectors created by the two possible headings and corrupted by measurement error—if a particular state of the world exists. Let $\langle \mathbf{AB} \rangle$ represent the hypothesis that observer motion **A** (e.g. simulated motion straight ahead) was presented in the first interval and observer motion **B** (e.g. simulated motion to the right of straight ahead) was presented second, and let $\langle \mathbf{BA} \rangle$ represent the hypothesis that **B** was presented first and **A** second. The likelihood ratio for a given observation (Φ_1, Φ_2) is given by: $L = P(\Phi_1, \Phi_2 | \langle \mathbf{AB} \rangle)/P(\Phi_1, \Phi_2 | \langle \mathbf{BA} \rangle)$. In the two-interval, forced-choice task, the observer maximizes percent correct by responding that the first interval contained the stimulus corresponding to straight ahead when the likelihood ratio is greater than 1 and by responding that the second interval contained the straight-ahead motion when the likelihood ratio is less than 1.

For some tasks, one can calculate the distributions of the likelihood ratio analytically given each of the two hypotheses and then directly calculate the discriminability of the two sets of signals and noises (e.g. Geisler, 1984). This proved too difficult for the work presented here, so a direct Monte Carlo simulation was used instead. Each experiment was simulated trial by trial. In each trial, both signals (the set of directions and speeds associated with each of the possible observer motions) were generated and corrupted by a set of random variables drawn from the appropriate noise distributions. The likelihood ratio was computed as described above, and the model selected the appropriate hypothesis depending on whether the likelihood ratio was greater or less than 1. The motion increment (2D direction or speed or 3D heading) was varied using a 2-down/1-up staircase procedure to find the ideal discriminator's 71%-correct threshold.

How is heading information distributed in the optic flow field?

One of the motivations for Gibson's conceptualization of the optic flow field was to provide a tool for determining where the best information was for various navigation tasks. For instance, he wanted to determine where the best information lay for visual guidance of an aircraft landing (Gibson, Olum & Rosenblatt, 1955). Gibson's approach, however, did not yield a means for quantifying the information contained in various parts of a given optic flow field.

The approach outlined here provides a means for quantification for a wide variety of observer motions and scenes. Figure 8 displays pairs of flow fields created by observer motion relative to a frontoparallel plane. In (A) the two flow fields were created by translation 1 deg to the left (open arrows) and 1 deg to the right (solid arrows) of straight ahead. In (B) the headings are 4 (open) and 6 deg (solid) to the right, and in (C) the headings are 50 and 90 deg (parallel to the plane) to the right. The changes in heading lead to changes in the vector lengths (speeds) and directions. Most

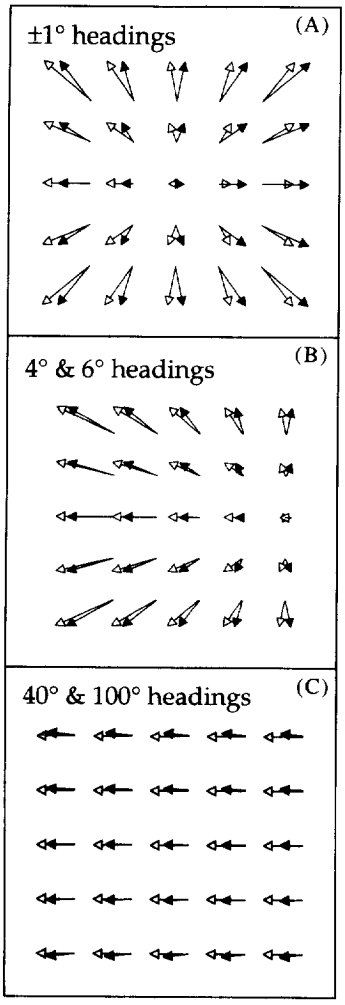


FIGURE 8. Flow fields created by two pairs of headings relative to a frontoparallel plane. In all cases, the field of view is 10 deg square. (A) Headings are 1 deg to the left and 1 deg to the right of straight ahead. The flow field associated with the former heading is indicated by the solid arrows and the one associated with the latter heading is indicated by the open-headed arrows. (B) Headings are 4 deg (solid arrows) and 6 deg (open-headed arrows) to the right. (C) Headings are 40 deg (solid arrows) and 100 deg (open-headed arrows) to the right. In all three panels the differences in directions and magnitudes between the two flow fields vary across the stimulus field. This variation exemplifies how the amount of information for discriminating the two headings varies from one stimulus region to another. The magnitudes of the differences are smaller in (C) even though the difference in headings is 30 times greater than in (A) and (B).

importantly, the magnitudes of those changes vary across the stimulus field. For example, in Fig. 8(A) the differences in vector directions are large directly above and below the headings, but go to zero on either side; the opposite is true for the differences in vector magnitudes. In signal-detection terms, the amount of information available for discriminating two signals is just the difference between the signals expressed in units of the parameters (σ_s and σ_d) of the measurement error distributions. We can show how useful different parts of the field are for discriminating these pairs of headings by plotting the differences in directions and speeds between the two flow fields in each pair divided by the appropriate standard deviations.

Figure 9 plots the differences in vector directions and speeds for the pairs of flow fields depicted in Fig. 8. Brightness is proportional to the magnitudes of the differences divided by the appropriate noise standard deviation computed from the curves in Fig. 6, so brighter areas represent regions in the stimulus field with larger differences relative to the noise. Stated another way, the brightest areas represent the regions in the stimulus field that contain the greatest information for discriminating the two headings. The top-left panel of Fig. 9 shows that the largest differences in vector directions occur above and below the center of the stimulus field when the two alternatives are nearly straight ahead; the top-right panel shows that the largest speed differences occur to the left and right of center. The bottom-left panel of Fig. 9 shows that the largest differences in direction occur in the upper- and lower-right corners of the field when the two alternatives are off to the right; the bottom-right panel shows that the speed differences increase monotonically from left to right. The two middle panels show that when the heading is towards the edge of the display, the pattern of differences is intermediate between the other two types. Clearly, the regions of the visual field containing the largest changes in flow vector directions and speeds differ depending on the headings from which the observer must choose.

Figures 10 and 11 show the application of the ideal observer model to the task of discriminating the alternative headings in Fig. 8. These figures were constructed by forcing the ideal observer to make the discriminations using only information in small regions of the visual field. Brightness is proportional to the percentage of correct responses the model attained while using only the information available at that position in the field. These are the results of Monte Carlo simulations, so the plots are somewhat noisy. The noises built into the ideal observer vary with retinal eccentricity and this property is included in the demonstration figures.

In Fig. 10, the model is "fixating" the center of the stimulus field. The three panels display the model's performance at distinguishing headings similar to those represented in the corresponding panels of Fig. 8. The differences between the headings used in this figure are smaller; the actual headings are 0.25 deg to the right and left of straight ahead [Fig. 10(A)], 4.25 and 5.75 deg to the right [Fig. 10(B)], and 50 and 90 deg to the right [Fig. 10(C)]. Besides the smaller difference between the two headings, the situation differed from that described in Fig. 8 in that the scene was a random 3D cloud of dots (depths from 70 to 570 cm) instead of a frontoparallel plane. As previously discussed, a random cloud provides less speed information than a plane.

As one might expect, the panels in this figure are quite similar to the corresponding panels in the left column of Fig. 9 indicating that the ideal observer's performance in this task is largely determined by direction rather than speed information. When the focus of expansion is visible [Fig. 10(A, B)], the most informative vectors are those above and below it; when the focus is not visible [Fig. 10(C)], the most informative vectors are at the top

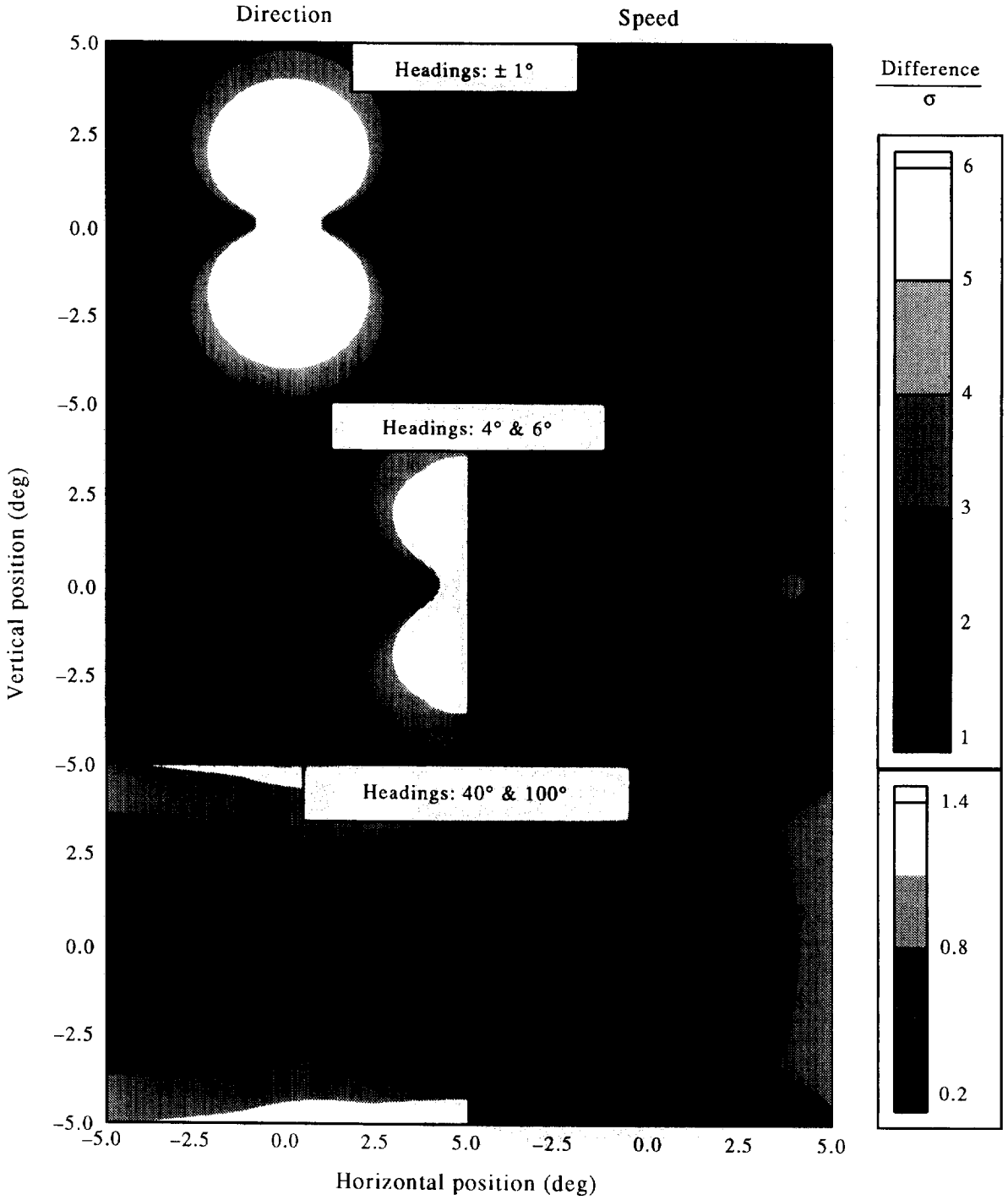


FIGURE 9. Spatial distribution of differences in vector directions and speeds for the flow fields displayed in Fig. 8. The top, middle, and bottom panels are derived from the flow field pairs in the corresponding panels of Fig. 8. In each case, brightness is proportional to the magnitude of the differences divided by the noise standard deviation (computed from Fig. 6). Brighter grays indicate larger differences relative to the noise and hence indicate better information for discrimination. The upper two rows of panels were rendered using the same brightness scale, but the scale in the bottom row of panels was magnified in order to make the variations more visible. The left panels display the differences in vector directions for various positions. The brightness peaks are actually much sharper than shown; they have been clipped to render the dimmer regions more visible. Scaled differences in vector directions are largest above and below the focus of expansion except in the bottom panel where the focus of expansion is not visible. The right panels display the differences in vector speeds. Differences in speeds are largest to the left and right of the focus of expansion except in the bottom panels where the focus of expansion is not visible.

and bottom of the field. These results are specific to a particular type of stimulus (high-contrast dots), scene geometry (a uniform 3D cloud), and task (discrimination of headings in the horizontal plane). Obviously, different outcomes would be obtained for different situations.

Figure 11 shows the effect of retinal eccentricity on the information in a radial flow pattern. Figure 11(A) is a

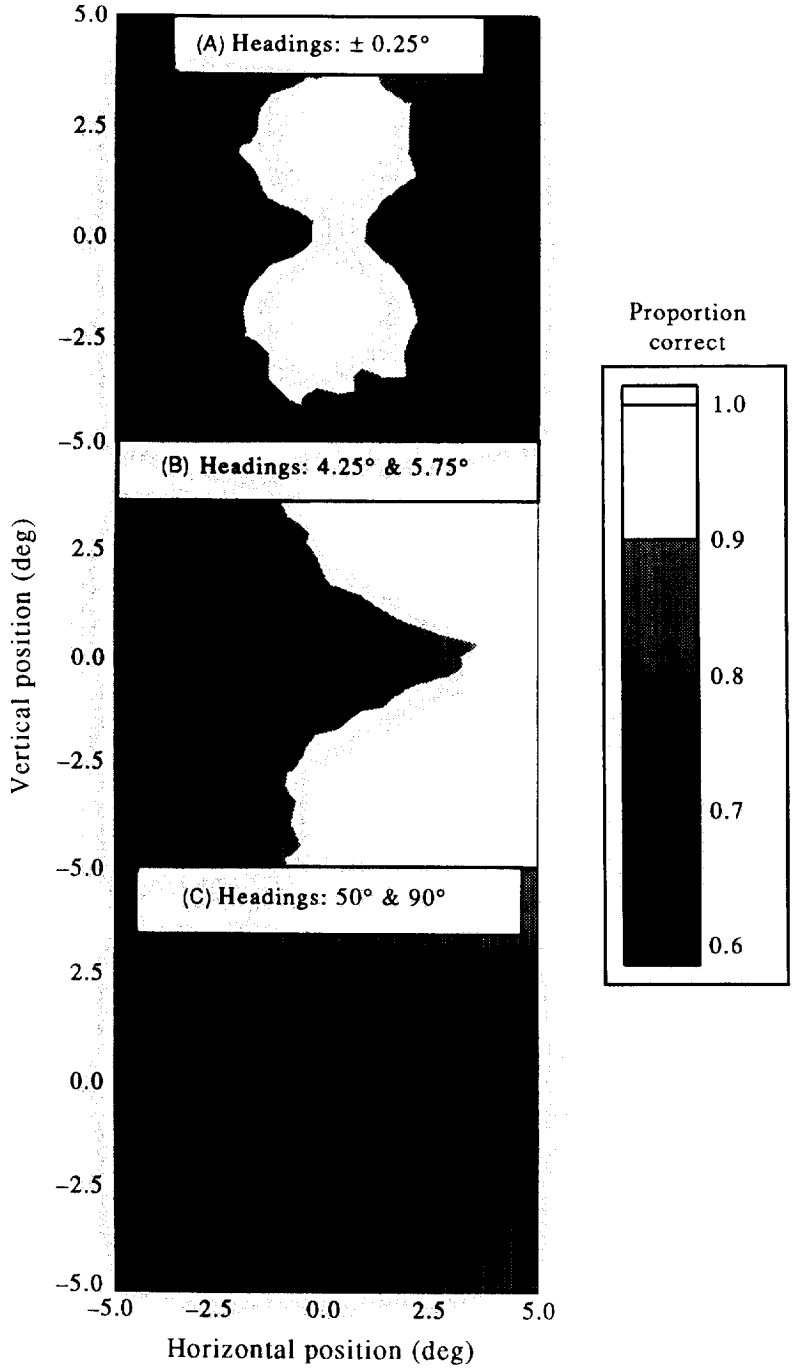


FIGURE 10. Model's performance as a function of location in the visual field for discriminating headings similar to those in Fig. 8. The actual headings were ±0.25 deg (A), 4.25 deg and 5.75 deg to the right (B), and 50 deg and 90 deg to the right (C). The stimulus was a random 3D cloud instead of a frontoparallel plane. The gray level at each point indicates the proportion of correct responses the model achieved using only the information at that position in the stimulus; the scale has been truncated at 60% correct performance (although chance performance was 50% correct) to increase the differences between the gray levels. Note the similarity of the panels to the corresponding left panels in Fig. 9, which shows that the ideal observer's performance is determined primarily by the information in flow vector directions.

copy of the top panel of Fig. 10; the headings are again 0.25 deg to the left and to the right of straight ahead. Figure 11(B) shows the model's performance at discriminating the same two headings with its peripheral "retina"; the center of the display is located 40 deg from its "fovea". The values above and below the focus of expansion are dimmer in this panel, which indicates that the information near the focus is reduced by peripheral viewing. Figure 11(C) illustrates this point in detail; it shows a vertical slice taken through the center of each of Fig. 11(A, B). The largest effect of retinal eccentricity occurs in the vicinity of the focus of expansion; the gap between the two curves representing central and peripheral performance is greatest at the center of the graph and decreases towards the sides. This reflects the fact that central and peripheral sensitivity to motion differs most for slow speeds which are observed around the focus of expansion.

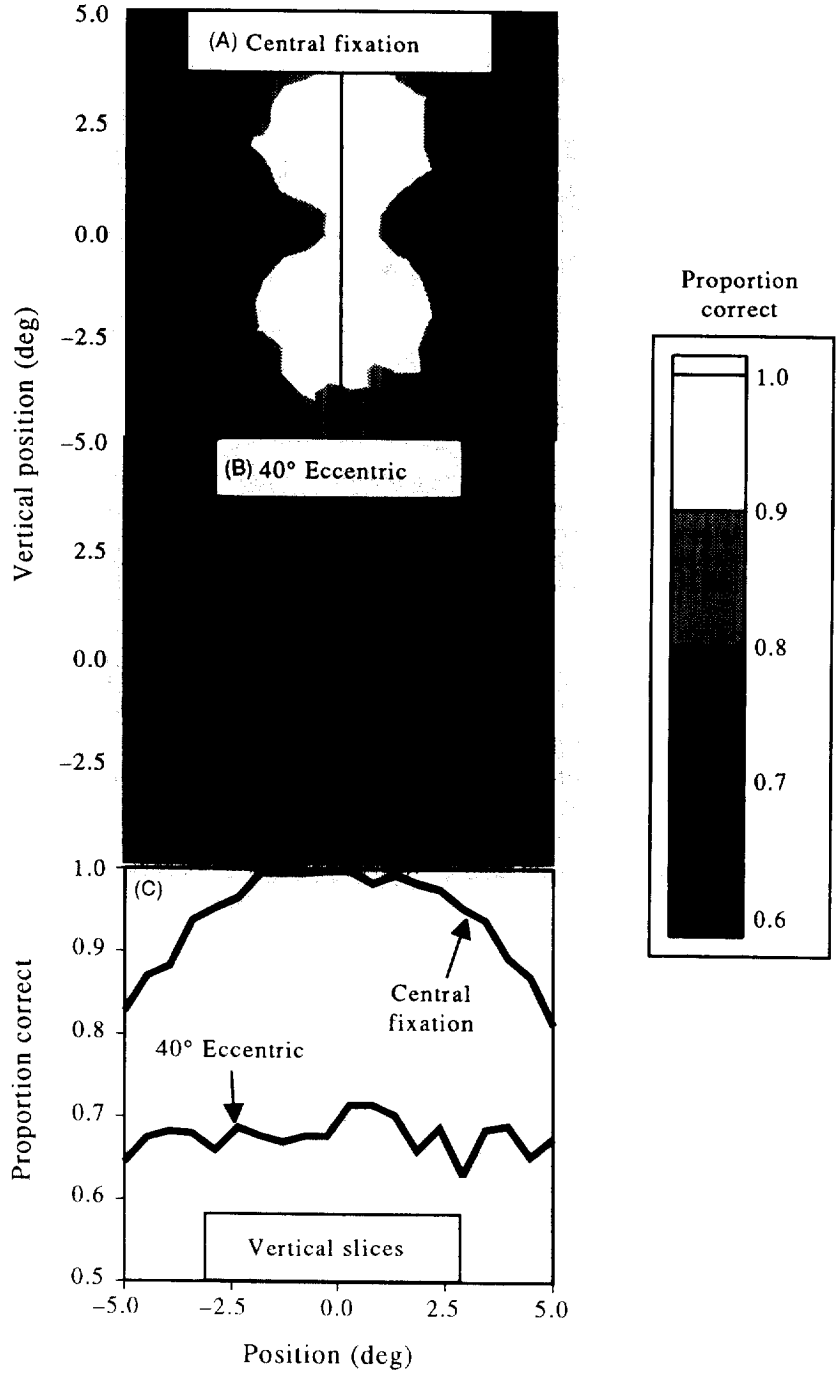


FIGURE 11. Model's performance as a function of location in the visual field for discriminating headings presented on different parts of the retina. The headings were ± 0.25 deg, which are the same as those in Fig. 10(A). The gray level at each point indicates the proportion of correct responses the model achieved using only the information at that position in the stimulus; the scale has been truncated at 60% correct performance (although chance performance was 50% correct) to increase the differences between the gray levels. (A) Performance when the stimulus patch was presented in the central visual field (retinal eccentricity = 0 deg). (B) Performance when it was presented in the periphery (retinal eccentricity = 40 deg). (C) Proportion correct as a function of vertical position in the stimulus for the positions indicated by the thin vertical lines in (A) and (B). The model's performance is noticeably reduced when the stimulus is presented in the periphery.

For discrimination of headings in the horizontal plane, these modeling outcomes demonstrate that: (1) the discrimination information (both directional and speed differences) is greatest near the focus of expansion and decreases monotonically away from it; (2) at a fixed distance from the focus of expansion, the information in flow directions is greatest directly above and below the focus of expansion and decreases monotonically to zero at the sides; and (3) at a fixed distance from the focus of expansion, the information in flow speeds is greatest directly to the left and right of the focus of expansion and decreases monotonically to zero above and below.

This application of the ideal observer should prove to be useful in the analysis of psychophysical results. It

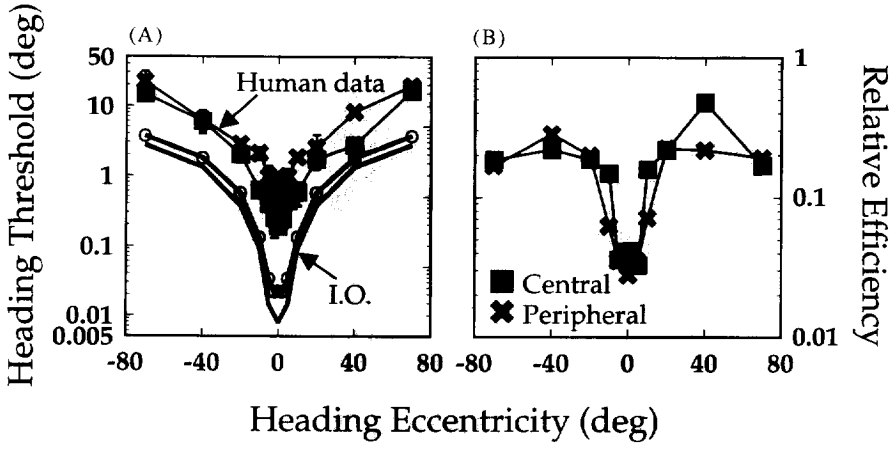


FIGURE 12. Human (observer JAC) and ideal discrimination thresholds as a function of heading eccentricity. (A) Human and model heading discrimination thresholds. Large symbols represent human thresholds at different retinal eccentricities: ■ central fixation; × 40 deg nasal retina. The curves represent the ideal observer's thresholds at the same retinal eccentricities: the lower curve for 0 deg and the upper curve (with ○) for 40 deg. (B) The relative efficiencies of the human observer; efficiency is defined as ideal threshold divided by human threshold. Symbols as in (A).

should also be useful in some applied settings such as the operation of vehicles. Specifically, if one can define the scene and the set of possible observer motions quantitatively, then the ideal observer described here can be used to calculate the distribution of heading information for that situation. This could prove useful for designing better lighting systems and for instructing vehicle operators.

Human efficiency in discriminating heading

Two effects are revealed in the heading discrimination data of Fig. 3. First, discrimination threshold rises by roughly a factor of 100 as heading eccentricity is increased from 0 to 70 deg. Second, although there is generally little effect of eccentric viewing, discrimination threshold with radial flow patterns is roughly three times lower with foveal as opposed to peripheral viewing. As mentioned earlier, both of these effects can be interpreted in two ways. In regard to the heading eccentricity effect, the advantage with radial flow could reflect a neural specialization for processing such a flow pattern, or it could be a consequence of the variation in discrimination information as the distance between the focus of expansion and the center of the stimulus is varied. If the former hypothesis is correct, then efficiency (defined as the ratio of human/ideal threshold) should be higher for radial than for laminar flow patterns. If the latter is correct, efficiency as a function of heading eccentricity should be constant. In regard to the retinal eccentricity effect with radial flow patterns, the foveal advantage may be a manifestation of special mechanisms for processing radial flow in the central visual field, or it could be a consequence of having receptive fields tuned to slower speeds in or near the fovea. If the former hypothesis is correct, efficiency should be greater for foveal than for peripheral viewing of radial flow patterns

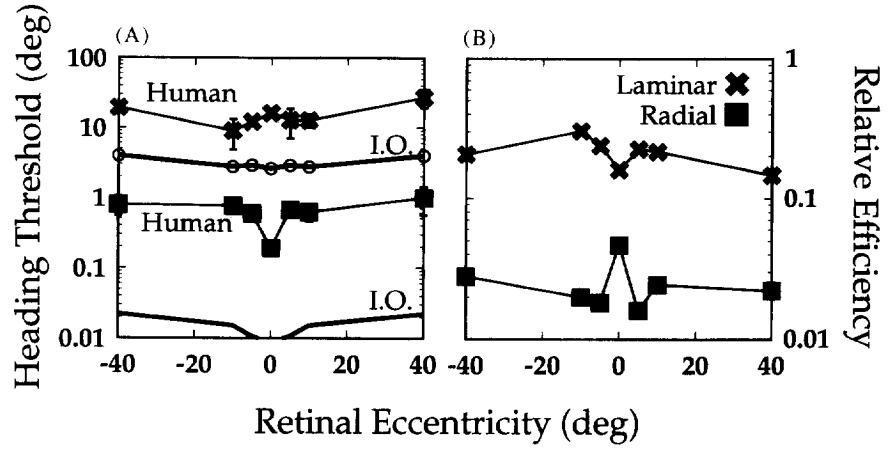


FIGURE 13. Human (observer JAC) and ideal discrimination thresholds as a function of retinal eccentricity. (A) Human and model heading discrimination thresholds. Large symbols represent human thresholds at different heading eccentricities: ■ 0 deg (radial flow); × 70 deg (mostly laminar flow). The two curves represent the ideal observer's thresholds: the lower curve for radial and the upper curve (with ○) for laminar flow. (B) The relative efficiencies of the human observer. Symbols as in (A).

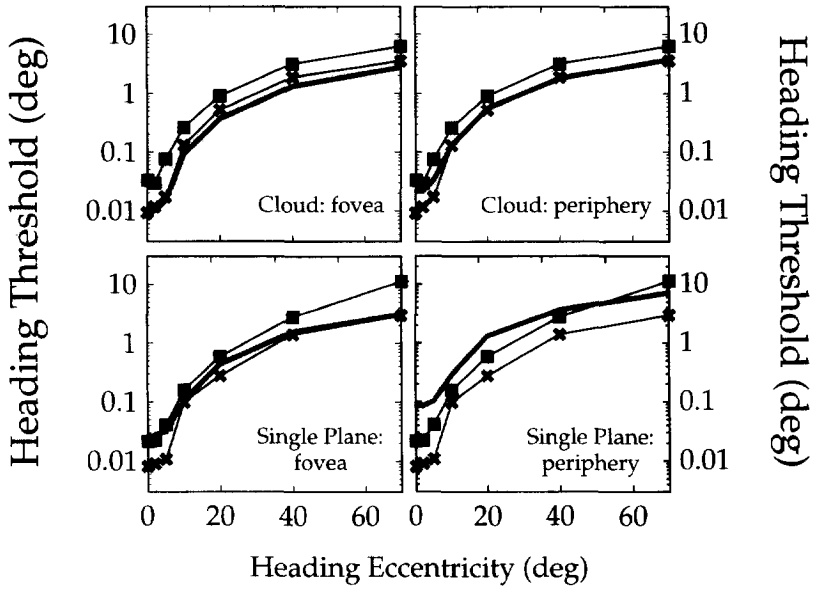


FIGURE 14. Comparison of ideal observer and least-squares algorithm of Koenderink and van Doorn (1987). Heading discrimination thresholds are plotted as a function of heading eccentricity. (A, B) Thresholds when the stimulus is a random 3-D cloud of dots. (C, D) Thresholds when the stimulus is a frontal plane (initial distance = 1050 cm). (A, C) Results for central viewing. (B, D) Results for peripheral (40 deg) viewing. In each panel, the thresholds for our ideal observer are represented by the bold lines, thresholds for Koenderink and van Doorn's least-squares algorithm are represented by ■, and thresholds for an ideal discriminator with the noise distributions assumed by Koenderink and van Doorn are represented by ×. (A, B) Model and human thresholds for the experiment of Crowell and Banks (1993). (A) Contains foveal data, (B) shows results for the far periphery (the other two models' functions are the same in both panels). The two ideal discriminators' thresholds are virtually identical for all heading eccentricities in the fovea and in the periphery for all but the smallest heading eccentricities. The least-squares algorithms thresholds are higher by a factor of 2–3. (C, D) Predictions of the three models for a hypothetical experiment in which the 3D cloud is replaced with a single plane at the same initial depth as the back of the cloud (1050 cm). In this case, our ideal observer predicts a slightly shallower heading eccentricity effect, and the overall level of performance predicted is quite different in the periphery.

and, if the latter is correct, efficiency should be constant across retinal eccentricity. Figures 12 and 13 show the results of this analysis for the heading and retinal eccentricity effects, respectively.

Figure 12(A) displays human and ideal heading discrimination thresholds as a function of heading eccentricity for foveal and peripheral viewing (0 and 40 deg). The flow fields presented to the ideal observer were identical in extent and in the number and distribution of dots to those presented to the human observers. The symbols are the human observer's data and the solid curves represent the model's thresholds. Figure 12(B) displays the relative efficiencies of the human observer.

The results of this analysis are interesting and counterintuitive. First, except for heading eccentricities near 0 deg, efficiencies are reasonably constant as a function of heading eccentricity. This means that much of the 2-log-unit degradation in discrimination performance with laminar as opposed to radial flow patterns is a consequence of changes in the discrimination information provided by the stimulus. Second, there is a decrease in efficiency around 0 deg which corresponds to those stimuli in which the focus of expansion was visible. This implies that the human observer is actually *less* efficient at extracting information from radial flow patterns than from other types. Of course, human observers are more *sensitive* to the information in radial flows, but this is a consequence of the richness of discrimination information provided by such flow patterns rather than a manifestation of more efficient use of the information. Third, the two efficiency curves are quite similar in shape indicating no large differences in efficiency between central and far peripheral viewing for any kind of flow pattern. This leads us to conclude that the observed differences in heading discrimination between central and far peripheral vision can be ascribed to differences in the front-end properties of the visual system that lead to more precise sensing of retinal image motion in the central field (Figs 15 and 16).

Figure 13 displays heading discrimination thresholds as a function of retinal eccentricity for two heading eccentricities (0 and 70 deg) and Fig. 13(B) plots relative efficiencies. The separation between the efficiency curves for radial and laminar flow is further evidence that human observers are less efficient in estimating heading from radial than from laminar flow patterns. The finer scale of retinal eccentricity in this figure as compared with Fig. 12 reveals that relative efficiency does in fact vary with retinal eccentricity although the variation, particularly with laminar flow patterns, is fairly small. The small increase in relative efficiency with radial flow patterns viewed in the fovea appears to be reliable.

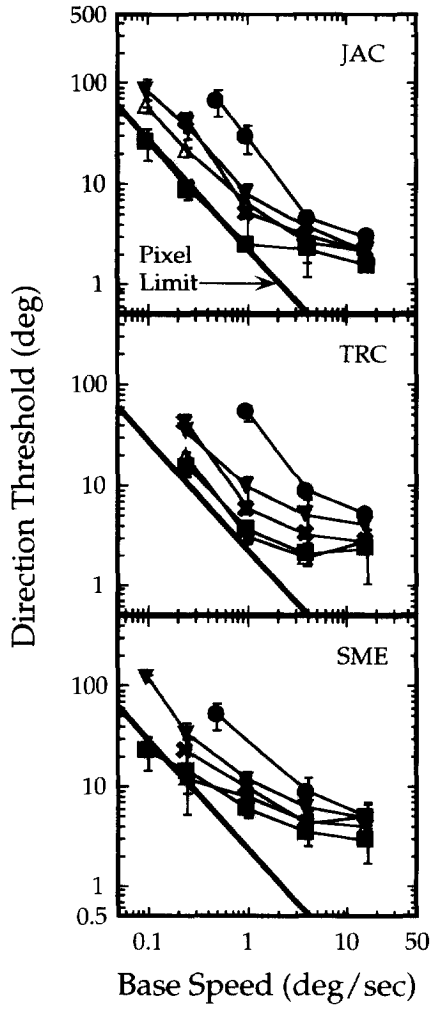


FIGURE 15. Single-dot direction discrimination thresholds plotted against base speed. The three panels display the data from the three observers. Different symbols represent different retinal eccentricities: ■ 0 deg, △ 2 deg; × 5 deg; ▼ 10 deg; ● 40 deg. The diagonal line shows the single-pixel limit as described in the text.

DISCUSSION

Comparison of results with previous reports

The approach presented here allows us to distinguish between two interpretations of psychophysical effects in heading perception. According to one interpretation, such effects reflect changes in the amount of information present in the stimulus (the optic flow field) after it has been measured by the human visual system. According to the other, the effects manifest differences in the efficiency with which the visual system computes a heading estimate from the measured flow field.

Effect of type of flow pattern. Crowell and Banks (1993a) observed a large degradation in the ability to discriminate headings with laminar as opposed to radial flow patterns. Regan and Beverley (1979, 1980) demonstrated the existence of channels sensitive to the changing size of an approaching stimulus ("looming"). Although they did not make such a prediction explicitly, this demonstration might lead one to expect greater efficiency in the estimation of heading from radial than from laminar flow patterns. On the contrary, the results of Fig. 12 show that the heading information contained in radial flow fields is used less efficiently than the information in laminar patterns. Thus, if channels sensitive to "looming" exist, they do not allow more efficient computation of heading than do channels sensitive to other types of motion.

Our analyses suggest the existence of two distinct regimes in regard to the type of flow pattern presented: when the focus of expansion is present in the stimulus, efficiencies are lower; when it is not present, efficiencies are higher. This effect is of roughly equal magnitude in the fovea and in the periphery. The fact that efficiencies are lower when the focus of expansion is visible seems to contradict Gibson's original notion of how heading is perceived. He argued that the focus of expansion provides a simple, salient marker for the direction of self-motion. Our analyses suggest that human observers are in a sense better (in terms of efficiency, not sensitivity) at locating the focus of expansion when it is *not* visible. The question is: "Does the lower efficiency with radial flow represent a performance limitation imposed by the stage of heading estimation, or is it a byproduct

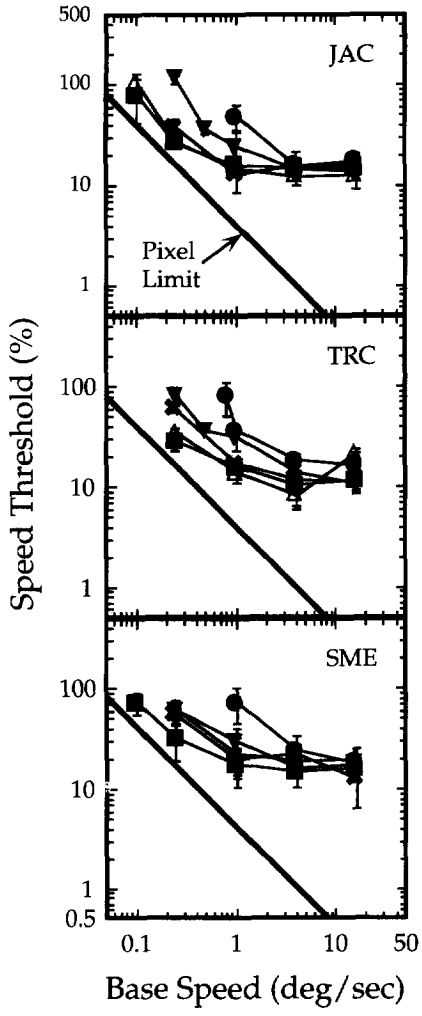


FIGURE 16. Single-dot speed discrimination thresholds plotted against base speed. The three panels display the data from the three observers. Different symbols represent different retinal eccentricities: ■ 0 deg, △ 2 deg; × 5 deg; ▼ 10 deg; ● 40 deg. The diagonal line shows the single-pixel limit as described in the text.

of not incorporating some property of the retinal motion estimation stage into the ideal discriminator model (see Fig. 4)?" An example of the second possibility would be an inability to resolve the directions of the dots near the focus of expansion because they change so rapidly over a small area. One could test this possibility by presenting the same stimuli with the area around and including the focus of the expansion masked off. Warren and Kurtz (1992) performed such an experiment and found a performance decrement when an area around the focus of expansion 10 deg or more in diameter was masked (they did not use smaller masks). For the purpose of testing the above hypothesis, we would need to know whether masking off a smaller area around the focus of expansion increased human observers' efficiency.

Effect of locus of retinal stimulation. The results depicted in Fig. 13 reveal little or no effect of retinal eccentricity. In particular, there were no significant differences in efficiency between central and peripheral vision, although there was a small gain in efficiency with radial flow fields in the fovea.

Warren and Kurtz (1992) have argued that heading estimates are more readily derived from radial flow in the central visual field than in the peripheral field. Our results are consistent with this claim, but the observed effect is very small (at most a factor of two).

There have been several claims in the literature that peripheral vision is better-suited to visual guidance of self-motion. For example, it has been hypothesized that peripheral vision is more influential than central in vection and postural control (e.g. Amblard & Carblanc, 1980; Brandt, Dichgans & Koenig, 1973; Held, Dichgans & Bauer, 1975). The mechanisms that compute heading from the internal representation of retinal image motion appear to be fairly homogeneous across the visual field. Thus, our results are inconsistent with the peripheral dominance hypothesis if we allow a generalization between our modeling of heading perception and the tasks involved in vection and postural control.

Overall efficiency. The relative efficiencies for the experimental conditions represented in Figs 12 and 13 ranged from 0.02 to 0.50. These efficiencies are rather high compared to those observed in many visual discrimination tasks, but there is an obvious explanation: the ideal observers that served as the performance benchmark in most analyses of visual discrimination (e.g. Banks, Geisler & Bennett, 1987; Barlow, 1958, 1962; Geisler, 1984, 1989) only incorporated stimulus properties, optical processing, and transduction among the photoreceptors whereas the observer presented here incorporated all visual processing stages up to and including the internal representation of dot motions. For this reason, the level of performance for the present ideal observer was lower than would have been the case for an observer with fewer stages incorporated, and estimated efficiencies were correspondingly higher.

It is important to consider what properties of the visual system limit the overall efficiencies. Stated another way, why is the observed discrimination performance poorer than exhibited by the model? There are several possibilities, some of which are discussed in the next section, but the primary cause of the observed gap is undoubtedly limited sampling efficiency (Legge, Kersten & Burgess, 1987). Specifically, the ideal observer uses the discrimination information contained in all of the dots and human observers probably do not. The evidence for this is as follows. One expects discrimination threshold of the ideal observer model to fall roughly in proportion to the square-root of the number of dots (assuming random assignment of dot position from one stimulus presentation to another). Discrimination thresholds in human observers fall, as expected, as the number of dots is increased from 2 to 10, but thresholds asymptote for larger numbers (Warren *et al.*, 1988). Relative efficiency should, therefore, decrease roughly as the square-root of dot number for displays containing more than 10–20 dots. The stimuli upon which our analysis is based (Crowell and Banks, 1993a) contained 500 dots on average, a range in which human observers almost certainly do not effectively utilize the information in each dot. One can estimate how much of the reduction in overall efficiency can be attributed to an inability to sample the information contained in the additional dots efficiently. Assuming that 10 dots are sampled efficiently and more are not, the square-root of the ratio of dots sampled divided by dots presented is $\sqrt{10/500}$, which is 0.14. One cannot place too much credence in this number because the experimental stimuli and task in Warren *et al.* (1988) were different from those in Crowell and Banks (1993b). Nonetheless, because the predicted efficiency of about 0.14 is similar the observed efficiencies, we speculate that much, but not all, of the reduced efficiency we observed is caused by inefficient use of the information in all the dots when more than 10 are presented.

Assumptions in ideal observer model

We had to make a number of assumptions in constructing the ideal observer. Here we review those assumptions and discuss their relevance in regard to our main conclusions.

Use of single-dot experiments to set noises. We used the results of single-dot discrimination experiments to construct the ideal observer for heading discrimination. In so doing, we assumed that the only important difference between the two types of discrimination is that heading requires an additional computation stage. However, an observer's performance in a discrimination experiment is also a function of his or her decision strategy. It is possible that our observers used either more or less efficient decision strategies in the single-dot experiment than in the heading experiment, which would lead to an error in the estimate of the noise associated with the measured optic flow that serves as input to the heading computation stage.

We believe that our observers used equally efficient decision strategies in the two types of tasks for the following reasons. First, observers were highly trained in

all of the experiments discussed here; they did not begin formal data collection until their thresholds asymptoted. It is unlikely, therefore, that the observers' strategies imposed a major limitation on performance in either single-dot or heading tasks. Second, human and ideal equivalent input noises in heading tasks are very similar. Crowell (1993) has compared human performance to that of our ideal observer in the presence of added stimulus noise. In such an experiment, performance typically is unaffected by added noise until a certain noise level is reached; this level is referred to as the "equivalent input noise" (Pelli, 1990), and constitutes an estimate of the internal noise associated with a given task. Recall that the internal noise in the ideal observer model is derived from human performance in the single-dot experiments; therefore, the similar equivalent input noises found by Crowell (1993) indicate that the single-dot results provide a reasonable estimate of the noise in the measurement of optic flow by the human visual system. Third, the efficiency of the human observer approaches 1 in the heading task when the number of dots is small (Crowell, 1993). That is to say, the heading discriminating thresholds of the human observer are nearly as low as those of the ideal observer when four or fewer dots are presented. Because the noises in the ideal observer were derived from the single-dot experiment, the similarity of thresholds argues against a significant difference in the efficiency of the decision strategy in the single-dot and heading experiments.

Single vector assumption. A vector with fixed direction and speed represented the motion of each dot during a stimulus presentation. The use of single vectors misrepresents the true dot motions because they actually accelerate over time; of course, their directions do not change. The magnitude of the acceleration depends on the ratio of the initial and final distances for a given dot and on the angle between the dot and the focus of expansion. The acceleration is largest for a heading of 0 deg and smallest for one of 90 deg. Fortunately, the accelerations were generally very small in the experiments considered here. For example, for headings of 0 deg, the average dot speed increased by only 5% over the course of a stimulus presentation. Thus, the single vector assumption is reasonable.

Independence assumption. We assumed that the noises associated with the measurement of the speed and direction of each dot were statistically independent and this assumption is probably incorrect. The assumption of independence leads to the best possible model performance because it allows errors to be averaged out across dots. If the noises were assumed to be perfectly correlated, there would be no advantage to incorporating more than one dot. If the independence assumption were correct, the accuracy of heading discrimination should increase in proportion to the square-root of the number of dots in the display. In contrast, Warren *et al.* (1988) and Crowell and Banks (1993b) have shown that discrimination performance asymptotes at 10–20 dots. A plausible explanation for this observation of no further improvement in performance with increasing dot number is the following. If the motions of dots in optic flow displays are analyzed by a limited number of sensors with spatially-extended receptive fields, then with a large numbers of dots, many are processed by the same sensor. If some of the noise associated with measuring dot motion is generated at the sensor, there will be a necessary spatial correlation among the noises associated with many of the dots. This would violate the independence assumption and limit performance with an increasing numbers of dots.

Despite the implausibility of the independence assumption, it is best at this point of our understanding to assume independent noises for the following reason. With the ideal observer approach, stimuli are deemed to provide a certain amount of information to the observer. The observer can either use all of the available information or (more likely) lose some of it to internal noise or to poor information-processing mechanisms and strategies. The observer is described in terms of the proportion of the available information that he or she can make use of across various transformations of the stimuli. The model of the observer is a set of mathematical assertions about aspects of the observer that cause the loss of certain kinds of information. As a corollary, one should leave out performance-degrading modeling steps unless there is clear evidence that such steps exist or one wishes to build in such steps and examine their effects in isolation. This is equivalent to assuming that all aspects of the observer for which one does not have specific evidence or a specific model are ideal (Watson, 1987). In the present context, we believe it is safer to make the independence assumption until there is more detailed information on the nature of correlations among motion measurements. In a forthcoming paper (Crowell & Banks, 1996), we will discuss results bearing on the independence assumption.

No bias assumption. We also assumed that the measurements of dot speed and direction are unbiased. This assumption may also be incorrect, particularly for the measurement of speed. Stone and Thompson (1992) have shown, for example, that the perceived speed of a grating moving at a constant speed is proportional to its contrast. In addition, the existence of phenomena such as motion contrast and induced motion (reviewed in Anstis, 1986) suggest that under certain conditions of relative motion the absolute motion of an object is not perceived veridically. We hasten to point out, however, that biases in the speed measurements are unlikely to have significant effects on discrimination tasks like those considered here. Such effects might be important if we were attempting to model perceived heading rather than just the precision of heading discrimination.

Use of speed information. The pattern of flow vector directions depends only on the position in the visual field relative to the focus of expansion, but the speed at a given point in the retinal image depends on the depth as well as on the heading (except at the focus of expansion, where the speed is zero). Because the ideal observer presented here made use of the speed differences, comparison of ideal and human performance is tantamount

to assuming that the human observer has knowledge of the distribution of depths in the scene. As stated above, we believe it is better at first to assume that this information is used until there is evidence to the contrary. Our results do not allow us to determine whether or not this assumption is correct; because we used a 3D random cloud of points, there was very little speed information available to begin with, and in fact the ideal observer performed almost identically when we restricted it to using only the information in the flow vector directions. Interestingly, Vishton, Nijhawan and Cutting (1994) have shown that the visual system takes advantage of independent depth information in some heading tasks. Specifically, they showed that the depth information conveyed by relative size and changing size improved the ability to estimate heading when the simulated observer motion included rotations.

Comparison with Koenderink and van Doorn's (1987) least-squares algorithm

Koenderink and van Doorn (1987) described and implemented a least-squares solution to the optic flow equations. This algorithm has been described as an ideal observer (van den Berg & Brenner, 1994). In this section, we will discuss the similarities and differences between their approach and ours.

As described above, our ideal observer has two main components: (1) a representation of the optic flow field (the signal) and the associated noise or error after measurement by the visual system; and (2) a statistically-optimal decision strategy for discriminating between pairs of flow fields created by different headings.

With regard to representing signals and noises, Koenderink and van Doorn (1987) took a similar approach to ours. That is, they represented the optic flow by a vector field and estimated the noise values from the results of psychophysical experiments on human velocity discrimination (McKee & Nakayama, 1984). The main difference lies in the amount of care taken in estimating the noise values. Koenderink and van Doorn assumed that each flow vector has associated with it a radially symmetric Normal noise distribution with a standard deviation equal to 10% of the vector magnitude; this is equivalent to assuming that (1) speed discrimination thresholds follow Weber's Law (with a Weber Fraction of 10%) for all base speeds; (2) direction discrimination thresholds are constant at about 5.7 deg for all base speeds; and (3) neither kind of discrimination depends on retinal eccentricity. As Fig. 6 shows, these assumptions are quite inaccurate in the peripheral retina except at high speeds.

Their approach also differs from ours in the second stage; in the context of Ideal Observer Theory, our approach is more appropriate for comparison with experiments such as those of Crowell and Banks (1993a). Koenderink and van Doorn (1987) assumed that the observer has no prior knowledge about the layout of the environment or about the set of motions that may be presented (rotational motions included). For the purposes of modeling our experiments, we have taken the opposite extreme; we have assumed that the observer knows exactly which two motions are going to be presented on any given trial and knows exactly the distribution of possible depths. If desired, we can calculate the effects of observer uncertainty about these parameters on performance, and we can perform experiments to estimate the degree of observer uncertainty. We believe it is better to assume complete certainty until we have a better idea of just how uncertain the observer is about the possible motions and scene geometry in experiments like ours.

Finally, given the constraints they assume, the Koenderink and van Doorn algorithm is not quite optimal. Their algorithm finds the set of motion parameters and the scene geometry that minimize the sum of the squared differences between expected and observed flow fields; an algorithm that minimizes the squared differences divided by the noise standard deviation for each flow vector performs better, although for the conditions of our experiments the improvement in performance is generally less than a factor of two.

Are these differences between the two approaches significant? The obvious way to answer this question is to compare the performance of the two algorithms for the same viewing conditions. Figure 14 shows heading discrimination thresholds for our ideal discriminator, for the least-squares estimator of Koenderink and van Doorn (1987),* and for an ideal discriminator that used the same noise distributions as Koenderink and van Doorn's least-squares algorithm. Our ideal observer's thresholds are indicated by the thick solid lines, the least-squares algorithm's by the squares, and the other ideal discriminator's by the crosses. Figure 14(A, B) shows the models' thresholds when the stimulus was a 3D cloud and Fig. 14(C, D) the thresholds when it was a plane. Figure 14(A, C) plots results for foveal viewing and Fig. 14(B, D) for peripheral viewing. The two ideal discriminators' thresholds are virtually identical in the fovea; they are quite similar in the periphery except at the smallest heading eccentricities where thresholds for our discriminator are higher. The thresholds of the Koenderink and van Doorn least-squares estimator are two to three times higher than those of the ideal discriminators because their algorithm does not weigh the errors appropriately according to the error distribution standard deviations and does not incorporate information about the possible motions and depths. Nonetheless, for the conditions of Crowell and Banks (1993a) experiment, the least-squares algorithm exhibits the heading eccentricity effect in much the same fashion as the ideal discriminators.

If we model experiments in which the dot speeds are slower, the algorithms behave less similarly. Figure 14(C, D) shows the discrimination thresholds of the three algorithms for a hypothetical experiment in the

*We used the same noise distributions and minimized the same error function that they did. We did not use their particular iterative algorithm for finding the minimum, but we were able to show that this departure had no discernible effect.

stimulus is a plane at the same initial depth (1050 cm) as the back of the 3D cloud in Crowell and Banks (1993a). In this case, our ideal observer exhibits a smaller effect of heading eccentricity (with foveal and peripheral viewing) and thresholds in the periphery are uniformly higher (the threshold for our ideal observer for radial flow is 10 times higher in the periphery than in the fovea). Thus, there are conditions in which the algorithms behave quite differently. Because the noises in our algorithm were derived from a broader range of experimental conditions, its performance is valid for a broader range of heading displays.

Applications and future directions

The ideal observer presented here allows us to quantify how the informativeness of the optic flow field varies with viewing conditions. This modeling tool can, therefore, be used to determine how the informativeness of the stimulus varies with several properties of interest, including the field of view, number of texture elements, geometry of the scene, and much more. As stated earlier, such a quantification allows one to measure human performance in terms of efficiency (the ratio of ideal/human threshold) and this in turn allows one to uncover those aspects of performance that reveal the properties of neural mechanisms and strategies human observers employ in these tasks. At this point, application of the ideal observer model is somewhat limited because it can only be applied to the discrimination of directions or speeds of observer translations and to flow fields consisting of random dots. We have begun to expand the model's scope to incorporate rotational components of observer motion (Crowell & Banks, 1994). In the future, it would be quite useful to include other types of texture elements besides random dots (e.g. Flach *et al.*, 1992). In its current state, there are nonetheless several potentially useful applications of this modeling tool.

We showed in Figs 10 and 11 that one can calculate the spatial distribution of heading discrimination information for the set of stimuli used by Crowell and Banks (1993a). For those stimuli, the distribution proved to be rather simple. The discrimination information was greatest near the focus of expansion (when it was visible) and fell monotonically with increasing distance from that point. The information contained in vector directions decreased most rapidly along an axis passing through the two possible headings, whereas the information in vector speeds decreased most rapidly along the orthogonal axis. Naturally, the spatial distribution of heading information will not always be so simple because it depends on the direction of self-motion relative to the stimulus (that is, the heading eccentricity), the scene geometry, and more. An important direction for future work will be to apply this modeling tool to the results of other previous studies of heading perception and to real-life situations. As an example of the former, consider the studies on maintaining altitude in flight simulations by using different types of texture (e.g. Flach *et al.*, 1992). Expanding the ideal observer model to allow the incorporation of lines, one could determine whether the superior performance of human observers when presented lines parallel to the forward motion is a consequence of greater informativeness of such lines or the use of more efficient mechanisms or strategies on the observers' part. As an example of the latter, one could use our approach to determine where in a pilot's visual field the most informative changes in flow would occur if the aircraft's heading deviated horizontally or vertically from the desired flight path.

CONCLUSION

We developed an ideal observer for the discrimination of heading from random-dot flow fields. It can be used to quantify the discrimination information available for any translation through any scene represented by dots. Using this ideal observer, we showed that the spatial distribution of discrimination information varies significantly depending on the viewing situation, particularly on the direction of simulated observer translation relative to the stimulus patch (the heading eccentricity). This sort of analysis can be used in a wide variety of situations including applied settings.

We compared human and ideal observer performance in discriminating headings with different patterns of flow presented on different parts of the retina. Efficiency was reasonably constant for different flow patterns and quite constant for different retinal eccentricities. This outcome shows that most of the variation in the ability to estimate heading from the flow patterns and retinal loci considered here is due to changes in the discrimination information provided by the stimulus after measurement by early stages of visual processing.

REFERENCES

Adelson, E. H. & Bergen, J. R. (1985). Spatiotemporal energy models for the perception of motion. *Journal of the Optical Society of America A, 2,* 284–299.

Albright, T. D. (1984). Direction and orientation selectivity of neurons in visual area MT of the macaque. *Journal of Neurophysiology, 52,* 1106–1130.

Amblard, A. & Carblanc, B. (1980). Role of foveal and peripheral visual information in the maintenance of postural equilibrium in man. *Perceptual & Motor Skills, 51,* 903–912.

Andersen, G. J. & Braunstein, M. L. (1985). Induced self-motion in central vision. *Journal of Experimental Psychology: Human Perception & Performance, 11,* 122–132.

Anstis, S. (1986). Motion perception in the frontal plane. In Boff, K. R., Kaufman, L. & Thomas, J. P. (Eds), *Handbook of perception and human performance, Vol. 1: Sensory processes and perception.* New York: Wiley.

Banks, M. S., Geisler, W. S. & Bennett, P. J. (1987). The physical limits of grating visibility. *Vision Research, 27,* 1915–1924.

Barlow, H. B. (1958). Temporal and spatial summation in human vision at different background intensities. *Journal of Physiology, 141,* 337–350.

Barlow, H. B. (1962). A method of determining the overall quantum efficiency of visual discriminations. *Journal of Physiology, 160,* 155–168.

van den Berg, A. V. & Brenner, V. (1994). Humans combine the optic flow with static depth cues for robust perception of heading. *Vision Research, 34(16),* 2153–2167.

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

Crowell, J. A. (1993). On the efficiency of heading judgments. University of California at Berkeley: unpublished doctoral dissertation.

Crowell, J. A. & Banks, M. S. (1993a). Perceiving heading with different retinal regions and types of optic flow. *Perception & Psychophysics, 53*, 325–337.

Crowell, J. A. & Banks, M. S. (1993b). Relative efficiency of 2D and 3D motion judgments. *Investigative Ophthalmology & Visual Science (Supplement), 34*, 3269.

Crowell, J. A. & Banks, M. S. (1994). An ideal observer for heading judgments with rotations. *Investigative Ophthalmology & Visual Science (Supplement), 35(4)*, 2000.

Crowell, J. A. & Banks, M. S. (1996). Sources of inefficiency in heading perception. Manuscript in preparation.

De Bruyn, B. & Orban, G. A. (1988). Human velocity and direction discrimination measured with random dot patterns. *Vision Research, 28*, 1323–1335.

Flach, J. M., Hagen, B. A. & Larish, J. F. (1992). Active regulation of altitude as a function of optical texture. *Perception & Psychophysics, 51*, 557–568.

Fleet, D. J. & Jepson, A. D. (1989). Computation of normal velocity from local phase information. University of Toronto Technical Report on Research on Biological and Computational Vision RBCV-TR-89-27.

Geisler, W. S. (1984). Physical limits of acuity and hyperacuity. *Journal of the Optical Society of America A, 1*, 775–782.

Geisler W. S. (1989). Sequential ideal-observer analysis of visual discriminations. *Psychological Review, 96*, 267–314.

Gibson, J. J. (1950). *Perception of the visual world.* Boston: Houghton–Mifflin.

Gibson, J. J. (1966). *The senses considered as perceptual systems.* Boston: Houghton–Mifflin.

Gibson, J. J., Olum, P. & Rosenblatt, F. (1955). Parallax and perspective during aircraft landings. *American Journal of Psychology, 68*, 372–385.

Green, D. M. & Swets, J. A. (1966). *Signal detection theory and psychophysics.* Huntington, N.Y.: Robert E. Krieger.

Heeger, D. J. (1987). Model for the extraction of image flow. *Journal of the Optical Society of America A, 4*, 1455–1471.

Held, R., Dichgans, J. & Bauer, J. (1975). Characteristics of moving visual scenes influencing spatial orientation. *Vision Research, 15*, 357–365.

Howard, I. P. & Heckmann, T. (1989). Circular vection as a function of the relative sizes, distances, and positions of two competing visual displays. *Perception, 18*, 657–665.

Koenderink, J. J. & van Doorn, A. J. (1987). Facts on optic flow. *Biological Cybernetics, 56*, 247–254.

Koenderink, J. J., van Doorn, A. J. & van de Grind, W. A. (1985). Spatial and temporal parameters of motion detection in the peripheral visual field. *Journal of the Optical Society of America A, 2*, 252–259.

Legge, G. E., Kersten, D. & Burgess, A. E. (1987). Contrast discrimination in noise. *Journal of the Optical Society of America A, 4*, 391–404.

McKee, S. P. & Nakayama, K. (1984). The detection of motion in the peripheral visual field. *Vision Research, 24*, 491–500.

Pelli, D. (1990). The quantum efficiency of vision: In Blakemore, C. (Ed), *Vision: Coding and efficiency*. Cambridge: Cambridge University Press.

Press, W. H., Flannery, B. P., Teukolsky, S. A. & Vetterling, W. T. (1988). *Numerical recipes in C.* Cambridge: Cambridge University Press.

Regan, D. & Beverley, K. I. (1979). Visually guided locomotion: Psychophysical evidence for a neural mechanism sensitive to flow patterns. *Science, 205*, 311–313.

Regan, D. & Beverley, K. I. (1980). Visual responses to changing size and to sideways motion for different directions of motion in depth: Linearization of visual responses. *Journal of the Optical Society of America, 70*, 1289–1296.

Rieger, J. H. & Toet, L. (1985). Human visual navigation in the presence of 3D rotations. *Biological Cybernetics*, 377–381.

Rose, A. (1948). The sensitivity performance of the human eye on an absolute scale. *Journal of the Optical Society of America, 38*, 196–208.

Royden, C. S., Banks, M. S. & Crowell, J. A. (1992). The perception of heading during eye movements. *Nature, 360*, 583–585.

van Santen, J. P. H. & Sperling, G. (1985). Elaborated Reichardt detectors. *Journal of the Optical Society of America A, 2*, 300–321.

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

Vishton, P. M., Nijhawan, R. & Cutting, J. E. (1994). Moving observers utilize static depth cues in determining their direction of motion. *Investigative Ophthalmology & Visual Science (Supplement), 35(4)*, 2000.

Warren, R. (1982). *Optical transformation during movement: Review of the optical concomitants of egomotion* (NTIS Tech. Rep. No. AD-A122275). Columbus, Ohio: Ohio State University.

Warren, R. (1988). Visual perception in high-speed low altitude flight. *Aviation, Space, & Environmental Medicine, 59(Suppl. 11)*, A116–A124.

Warren, W. H. & Hannon, D. J. (1988). Direction of self-motion is perceived from optical flow. *Nature, 336*, 162–163.

Warren, W. H. & Hannon, D. J. (1990). Eye movements and optical flow. *Journal of the Optical Society of America A, 7*, 160–169.

Warren, W. H. & Kurt, K. J. (1992). The role of central and peripheral vision in perceiving the direction of self-motion. *Perception & Psychophysics, 51*, 443–454.

Warren, W. H., Morris, M. W. & Kalish, M. (1988). Perception of translational heading from optical flow. *Journal of Experimental Psychology: Human Perception & Performance, 14*, 646–660.

Watson, A. B. (1987). The ideal observer concept as a modeling tool. In *Frontiers of Visual Science: Proceedings of the 1985 Symposium* (pp. 32–37). Washinton, D.C.: National Academy of Sciences.

Watson, A. B. & Ahumada, A. J. (1985). Models of human visual-motion sensing. *Journal of the Optical Society of America A, 2*, 322–341.

Westheimer, G. & Wehrhahn, C. (1994). Discrimination of direction of motion in human vision: *Journal of Neurophysiology, 71*, 33–37.

Wolpert, L. (1987). Field of view versus retinal region in the perception of self-motion. Ohio State University: Unpublished doctoral dissertation.

Wolpert, L. (1988). The active control of altitude over differing texture. In *Proceedings of the 32nd Annual Meeting of the Human Factors Society*, 15–19.

Acknowledgements—We thank Kirk Swenson for software assistance and Rowan Candy, Sheryl Ehrlich, and Johanna Weber for their participation as observers. This research was supported in part by NSF Research Grant DBS-9309820 and NIH Research Grant HD-19927 to M. S. Banks. J. A. Crowell was supported by NIH Training Grant EY-07043.

APPENDIX

Single-dot Discrimination Experiments

McKee and Nakayama (1984) showed that base speed and retinal eccentricity affect speed discrimination thresholds. They reported that the Weber Fraction for speed discrimination is constant (5–10%) at high speeds, but, as the base speed is decreased, there comes a point at which the Weber Fraction begins to rise monotonically. This point occurs at higher base speeds in the peripheral visual field. We also know from the work of De Bruyn and Orban (1988) and Westheimer and Wehrhahn (1994) that direction discrimination is poorer at slow speeds. Taken together, these results suggest that measurement error in the flow field is larger relative to the flow vector magnitude for lower

speeds, particularly in the peripheral visual field. In the single-dot experiment, the measurement error functions for speed and direction of motion were mapped out using five retinal eccentricities and four to six speeds.

Observers. Three observers participated in these experiments. JAC was a 7-diopter myope, corrected to normal, with considerable experience in similar experimental tasks. TRC was emmetropic with some experience in similar tasks. SME was a 2-diopter astigmat, corrected to normal, with no prior experience in these tasks.

Stimuli. Experiments were run on an Apple Macintosh IIfx computer with a 14-inch Apple color monitor. Stimuli were updated at the 66-Hz frame rate. The stimuli consisted of a single dot moving across the CRT in each of two intervals. The dot was an illuminated pixel subtending 1.8 by 1.8 min at the viewing distance of 70 cm. The dot's luminance was 7.7 c/deg/m^2 (measured when the screen was flooded); this luminance was chosen to match the apparent brightness of the dots used by Crowell and Banks (1993a). Background luminance was 0.01 c/deg/m^2. The CRT provided the only light in the room. The base direction of the dot's motion was always 45 deg up and to the right.* The dot passed near the screen center at the mid-point of its trajectory, but starting position and duration contained random components to minimize positional cues. The random component of starting position was uniformly distributed in X and Y with a full width of 20–50% of the total trajectory length; the random component of duration was also uniform with a width of 150–300 msec (conditions with higher thresholds had larger random components). Mean stimulus duration was 500 msec.

Dot motion on a CRT is, of course, discrete rather than continuous, but the motion appeared reasonably continuous at all but the slowest speeds. Some of the foveal thresholds, however, were in fact limited by the pixel size; those thresholds will be pointed out in the Results section. The fact that a small number of the single-dot thresholds was limited by the display's properties is not problematic for our analysis because the same limitation existed in the heading experiments.

Procedure. Observers fixated a stationary LED that was positioned to set the retinal eccentricity of the center of the stimulus patch to 0, 2, 5, 10, or 40 deg in the nasal retina (none of the stimuli fell on the blind spot). Direction and speed discrimination thresholds were measured at base speeds of 0.25, 1, 4, and 16 deg/sec (JAC and SME also collected some data at lower speeds). Experimental runs were blocked by condition. A 2-temporal-interval, forced-choice psychophysical procedure was used to estimate discrimination threshold. On a given trial, two 500 msec dot motions were presented, separated by a 300 msec blank interval. Observers responded by pressing a key to indicate whether the motion in the second interval appeared to be faster (speed discrimination) or in a more clockwise direction (direction discrimination) than the motion in the first interval. Feedback was given after each response. The value of the direction or speed increment was varied using a 2-down, 1-up staircase, which converges to the 71% correct point. The staircase was terminated after 12 reversals and threshold was taken to be the mean of the last 10 reversals. Three to five runs were conducted in each condition and the resulting thresholds averaged.

Results

Single-dot direction and speed discrimination thresholds are plotted in Figs 15 and 16, respectively. The figures display thresholds as a function of base speed for the three observers; different symbols represent retinal eccentricities of 0, 2, 5, 10, and 40 deg.

As expected, direction and speed discrimination thresholds were roughly constant at high speeds and increased steeply at low speeds (McKee & Nakayama, 1984; DeBruyn & Orban, 1988; Westheimer & Wehrhahn, 1994). Also as expected, the speed below which thresholds began to rise was higher at greater retinal eccentricities (McKee & Nakayama, 1984). The speed discrimination thresholds were about twice as high as McKee and Nakayama's, but this disparity is undoubtedly due to our use of very small (1.8 min square) dots; they used lines several degrees long.

Presumably, direction and speed discrimination thresholds would have been lower at some of the slowest base speeds if the stimuli consisted of continuous motion. A few of the threshold values were close to the limit imposed by the pixel size of the CRT. For example, at the slowest base speed in the foveal speed discrimination condition (0.1 deg/sec), the dot moved only two pixels over the course of a trial; thus, the reported Weber fraction of $\sim 50\%$ corresponds to a difference of one pixel between the two intervals. The thick solid curves in Figs 15 and 16 represent the thresholds that would have been measured had the observers responded correctly whenever the dot's trajectory varied by a single pixel over the course of the two motion sequences. As mentioned earlier, this display limitation does not adversely affect our analysis because the same limitation applies to the stimuli in the heading experiments.

We did not observe any large, systematic variations between observers. The largest range of thresholds in any given condition was roughly a factor of two, and in most conditions, it was considerably smaller. Therefore, we chose to use thresholds averaged across the three observers to set the noises in the ideal observer.

*Pilot work revealed that discrimination thresholds varied by less than 10% as a function of base direction. The effects of interest here are much larger, so we used one base direction only in the main single-dot discrimination experiments.