
Perceiving heading in the presence of moving objects

William H Warren Jr, Jeffrey A Saunders

Department of Cognitive and Linguistic Sciences, Brown University, Providence, RI 02912, USA
Based on paper presented at the Conference on Binocular Stereopsis and Optic Flow, Toronto, Canada, 22–26 June 1993

Abstract. In most models of heading from optic flow a rigid environment is assumed, yet humans often navigate in the presence of independently moving objects. Simple spatial pooling of the flow field would yield systematic heading errors. Alternatively, moving objects could be segmented on the basis of relative motion, dynamic occlusion, or inconsistency with the global flow, and heading determined from the background flow. Displays simulated observer translation toward a frontal random-dot plane, with a 10 deg square moving independently in depth. The path of motion of the object was varied to create a secondary focus of expansion (FOE') 6 deg to the right or left of the actual heading point (FOE), which could bias the perceived heading. There was no effect when the FOE was visible, but when the object moved in front of it, perceived heading was biased toward the FOE' by $\sim 1.9^\circ$ with a transparent object, and $\sim 3.4^\circ$ with an opaque object. The results indicate that scene segmentation does not occur prior to heading estimation, which is consistent with spatial pooling weighted near the FOE. A simple template model based on large-field, center-weighted expansion units accounts for the data. This may actually represent an adaptive solution for navigation with respect to obstacles on the path ahead.

1 Introduction

In most models that recover heading from optic flow it is assumed that the flow field is produced by observer motion in a rigid environment (Bruss and Horn 1983; Heeger and Jepson 1992; Lappe and Rauschecker 1993; Longuet-Higgins and Prazdny 1980; Perrone 1992; Rieger and Lawton 1985; Tsai and Huang 1981; Waxman and Ullman 1985). Yet the world is populated by independently moving objects that violate this assumption, and people appear to navigate successfully on crowded sidewalks and busy freeways. In this paper, we attempt to determine whether human observers can accurately perceive their translational heading with respect to a stationary environment in the presence of a moving object. The answer appears to be a qualified yes—except when the moving object obscures the heading point. We present a simple template model based on large-field, center-weighted expansion units that accounts for the data. This may actually be an adaptive solution for navigation with respect to obstacles on the path ahead, rather than the frame of reference provided by the stationary surround.

When an observer translates through a rigid environment, a radial pattern of optic flow is generated with a focus of expansion (FOE) in the direction of heading (Gibson 1950). Recent experiments have confirmed that the visual system relies on this global radial flow pattern to perceive heading, pooling over local motion vectors throughout the field. Heading accuracies are within 1° – 2° of visual angle even with very low dot densities and considerable noise in local dot motions (Warren et al 1988, 1991). Further, vectors near the focus of expansion have been shown analytically and empirically to be more informative than those farther away (Crowell and Banks 1993; Koenderink and van Doorn 1987). Specifically, the FOE can be located by triangulating two or more vectors back to their common point of intersection; with noise in local motions, vectors farther from the focus will produce greater triangulation error.

These properties were exhibited by a simple neural model (Hatsopoulos and Warren 1991) inspired by primate visual areas MT and MST, in which a Widrow-Hoff training procedure resulted in a set of expansion templates with radially structured, center-weighted receptive fields (see also Lappe and Rauschecker 1993; Perrone 1992; Perrone and Stone 1994). The retinal flow pattern is often complicated further by an added component of observer rotation, such as a pursuit eye movement (Royden et al 1992; Warren and Hannon 1990).

An independently moving object generally adds a region of motion that is inconsistent with the radial structure of the flow pattern (figure 1). The exception is when the object moves toward the observer on a path parallel to the observer's path; only in this case is there a common focus of expansion and a rigid three-dimensional interpretation of the scene. The optic flow of the background is determined solely by observer motion, and its FOE specifies heading with respect to the stationary surround. The optic flow of the object is determined by both observer motion and object motion. One may think of the object as possessing a secondary focus of expansion (FOE'), which specifies the observer's instantaneous heading with respect to the moving object alone. It would clearly be advantageous to distinguish one's heading with respect to the background and the object and respond to them selectively.

There are several general hypotheses as to how the visual system might cope with moving objects:

- (a) *Spatial pooling*. One possibility is that moving objects are not explicitly segmented, but are treated as part of the global flow pattern. In effect, the visual system pools over all flow vectors to locate the FOE. This would lead to predictable errors in perceived heading. In figure 1, for example, the perceived heading would be somewhere in between FOE and FOE', depending on how their contributions were weighted.
- (b) *Segmentation → heading*. To distinguish observer motion from object motion, one could first segment regions that are likely to belong to the same surface, and then group those that are consistent with a common rigid three-dimensional motion (Adiv 1985). Background surfaces would thus be grouped together, yielding the observer-motion parameters, and each moving object would be grouped separately, yielding their motion parameters. Segmentation could be performed on the basis of relative motion, dynamic occlusion, or other surface information.
- (c) *Heading → segmentation*. An opposite approach first estimates the global observer-motion parameters and then segments discrepant regions; the heading estimate could be revised once these regions are removed. Observer motion may be estimated from optic flow together with depth, inertial, or positional information (Heeger and Hager 1988;

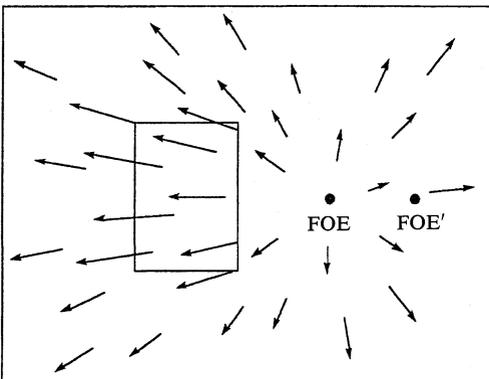


Figure 1. Schematic of flow field produced by translation toward a background plane with a moving object in the foreground. FOE is the focus of expansion for the background (the heading point), FOE' is the secondary focus of expansion for the object.

Thompson and Pong 1990; Zhang et al 1988), or from the flow alone (da Vitoria Lobo and Tsotsos 1991; Thompson et al 1993). Surface information such as dynamic occlusion could contribute to the segmentation.

(d) *Heading + segmentation*. Hildreth (1992) proposed an extension of Rieger and Lawton's (1985) differential-motion algorithm to recover heading and segment moving objects simultaneously. Any component of flow due to observer rotation is first eliminated by extracting the relative motion between neighboring elements at different depths. These difference vectors tend to radiate from the heading point, and they 'vote' for all candidate FOEs with which they are consistent. The FOE with the largest proportion of consistent vectors is taken to be the heading point, and coherent groups of inconsistent vectors are taken to indicate moving objects.

In the present experiments we examine whether human observers can indeed perceive their heading with respect to the background in the presence of moving objects, and attempt to differentiate several of the hypotheses. In this initial study, we tested the simple case of translation toward a frontal random-dot plane, with a single square object moving independently in depth. On the belief that performance is likely to be best in unconstrained conditions, we allowed observers to make free eye movements. For the case of a frontal plane, it is known that translational heading judgments are accurate under free fixation conditions (Warren et al 1988), and that the effects of rotation due to active eye movements can be discounted by the visual system (van den Berg 1992; Royden et al 1994; Warren and Hannon 1990).

We manipulated three main variables. First, the path of motion of the object was varied so that the FOE' appeared 6 deg to the right or left of the actual heading point. If the visual system were spatially pooling the flow, this would bias the perceived heading toward the FOE'. Second, to determine whether segmentation might be aided by information for object boundaries, we used three types of objects: opaque objects, which were defined by both relative motion and dynamic occlusion, transparent objects, defined only by relative motion between object dots and background dots, and black objects, defined only by dynamic occlusion with no relative dot motion. Note that there was more relative motion in the Transparent condition than the Opaque condition, because background dots were visible through the object. Third, the location of the moving object was varied with respect to the heading point. In experiment 1 they were on opposite sides of the screen, so the FOE was always visible, and in experiment 2 they were on the same side of the screen, so the FOE was obscured by the moving object. In experiment 3, Visible and Obscured trials were intermixed.

The predictions are as follows. (a) The spatial-pooling hypothesis predicts that both opaque and transparent objects should bias perceived heading toward the FOE'. The bias may be greater with opaque objects, because fewer background dots contribute to the perceived heading. (b) If, as proposed by Hildreth (1992), the visual system uses relative motion to segment or disregard the moving object, we would expect accurate performance with the transparent object. Performance with the opaque object could be worse because it contains less relative motion. (c) Alternatively, if the visual system uses dynamic occlusion to aid segmentation, we would expect accurate performance with the opaque object, owing to its enhanced boundary information, and possibly worse performance with the transparent object.

2 Experiment 1: visible focus of expansion

In the first experiment, the moving object was on the opposite side of the screen from the heading point, so the FOE was always visible. Three factors were manipulated, by means of the method of constant stimuli: the path of motion, transparency, and size of the object.

2.1 Method

2.1.1 Observers. Twelve students and staff at Brown University were paid to participate. All had normal or corrected-to-normal vision and, with the exception of William Warren, were participating in an optic-flow experiment for the first time. Four of these observers were removed because they performed at chance in ten practice trials that did not contain a moving object, leaving eight observers in the final group. Thus, our results can only be generalized to individuals who can perform the basic heading task reliably.

2.1.2 Displays. Displays were generated on a Silicon Graphics IRIS 4D/210 GTX workstation, and were presented on a raster monitor at 30 frames s^{-1} with a resolution of 1280 pixels horizontally \times 1024 pixels vertically and a 60 Hz refresh rate. Observers viewed the display binocularly with free fixation from a chin rest 43 cm from the screen. The screen was visible through a window in a matte black viewing box, and subtended a visual angle of 40 deg horizontally \times 32 deg vertically.

In all experiments, displays simulated observer translation toward a stationary background plane with a foreground object that moved independently in depth. The background consisted of 300 dots randomly positioned on a frontal plane, and the object was a frontal square with the same initial dot density as the background. In the Opaque condition, the object occluded the dots in the background. In the Transparent condition, the object dots were simply superimposed on the background, yielding an increase in density that was not noticeable without scrutiny. The object had an initial diameter of either 10 or 15 deg and contained 25 or 55 dots, respectively.

On each trial, the dots appeared for 1 s as a warning signal; this was followed by 1.5 s of motion, after which a 1 deg vertical probe line appeared; the probe and the last frame of dots remained visible until a response was made. The observers' task was to judge whether it appeared that they were heading to the left or right of the probe. The direction of observer translation was randomly varied between $\pm 0.5^\circ$, $\pm 1^\circ$, $\pm 2^\circ$, and $\pm 4^\circ$ on either side of the probe, which was defined as the heading angle, and the probe was randomly positioned at ± 4 , ± 6 , ± 8 , or ± 10 deg from the center of screen. Figure 2 shows a top-down view of the simulated environment and motion. In terms of dimensionless units of distance, the initial distance of the background plane and the moving object was 10 units. The relative speed between observer and background was 2 units s^{-1} (initial time to contact = 5 s) and between observer and object was 3 units s^{-1} (initial time to contact = 3.33 s). The center of the moving object

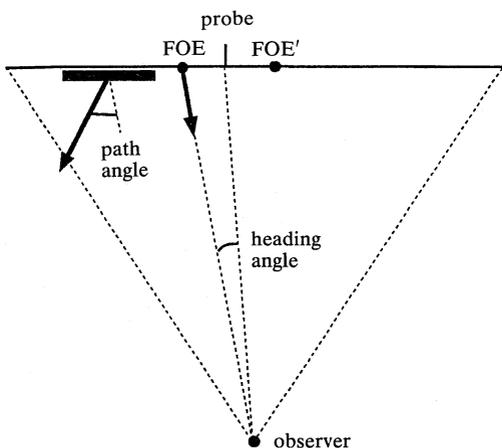


Figure 2. Top view of the display geometry (see text for details). (Relative motion between observer and background is represented as background motion.)

had an initial position of ± 6.0 deg relative to the center of the screen, constrained so that the object was always on the opposite side of the screen from the probe. Thus, the center of the object was initially 6 deg to 20 deg away from the heading point.

The difference angle between the path of motion of the object and the observer's path of motion, called the path angle, varied between -6° , 0° , and $+6^\circ$. It is equal to the visual angle between the FOE and the FOE', where positive values indicate that the FOE' is toward the center of the screen. For example, with a path angle of $+6^\circ$, the FOE' is 6 deg toward the center of the screen from the actual-heading point, and might bias perceived heading in that direction.

Trials were blocked by Transparent/Opaque condition, each block presented in a separate session in a counterbalanced order. All other variables were randomly varied. Each observer received a total of 768 test trials, providing 64 data points for each of the 12 heading estimates in the three-way design.

2.1.3 Procedure. Observers were asked to indicate with a keyboard press whether it looked as if they were heading to the left or right of the probe. They were told that a moving object would be present, and were instructed to ignore this object as much as possible and to base their responses on their perceived movement toward the background plane. At the beginning of the first session, there were 10 practice trials with feedback involving displays with no moving object, designed to familiarize the observers with the heading task. An additional 10 practice trials were presented without feedback at the beginning of each block, to familiarize observers with the object condition. There was no feedback on test trials.

2.1.4 Data analysis. The purpose in the experiments was to test whether a moving object systematically biases perceived heading, so measures of both the accuracy (constant error) and precision (variable error) of heading judgments were computed. The independent variable used in the analysis was the heading angle, the visual angle between the actual heading and the probe. To preserve the left/right symmetry of the task, the sign of the heading angle was chosen to be positive when the heading was toward the center of the screen relative to the probe. The dependent variable was then the percentage of 'center' responses. Analysis of practice trials indicated that if individual observers had a systematic bias in heading judgments without a moving object, it was toward or away from the center of the screen, rather than to the left or right.

For each observer and experimental condition, responses were combined across probe positions to yield the percentage of 'center' responses as a function of the heading angle. This function was fit by an ogive, and two parameters were extracted: the point of subjective equality (PSE) or heading angle at which observers performed at the 50% level, and the difference limen (DL) or visual angle between the PSE and 75%-correct level. The PSE can be interpreted as the perceived heading, and reversing its sign (owing to the definition of the heading angle) yields the constant heading error. The DL is comparable to a 75% reliability threshold, ie the precision in perceived heading about the PSE, a measure of variable error.

Four of the 96 heading estimates could not be fit accurately by an ogive (by the criterion $r > 0.8$) and were excluded from the analysis. In addition, 8 estimates with large PSEs ($|\text{PSE}| > 8^\circ$) were truncated to a value of $\pm 8^\circ$ because we were not confident in extrapolating far beyond the tested range, yielding a conservative estimate of the heading error.

2.2 Results and discussion

Under these conditions, the moving object had no effect on perceived heading. Figure 3 is a plot of the mean constant heading error as a function of path angle (the visual angle between FOE' and FOE) for each object type and object size. Three-way

multivariate repeated-measures ANOVAs revealed no significant main effects or interactions for either constant error or variable error. Thus, for those observers who can perform the basic heading task, the presence of a moving object does not affect their performance, at least when the FOE is visible.

There was an overall constant heading error of 1.25° , reflecting a slight bias toward the center of the screen across conditions. An analysis of practice trials with no moving object revealed a similar center bias, which suggests that it is not due to the presence of the moving object. In experiment 2, we added a block of control trials without a moving object to examine this more carefully. The mean DL was 1.81° , comparable to previously observed heading thresholds without moving objects.

It is possible that heading judgments are based on the location of the focus of expansion *per se*, or on the most informative vectors around it. Some observers reported a strategy of fixating the FOE during the trial, and then basing their response on the relative position of the probe when it appeared at the end of the trial. To test the region of the flow field near the FOE, in the next experiment we placed the moving object in front of the heading point on every trial.

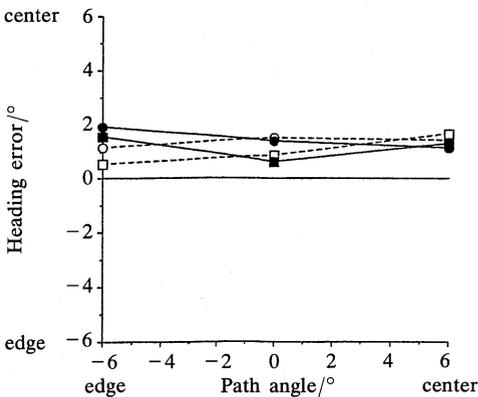


Figure 3. Mean heading error in experiment 1 (visible FOE). Path angle is equal to the visual angle between FOE and FOE'. Filled symbols represent opaque objects and open symbols transparent objects; circles and squares represent objects of diameter 10 deg and 15 deg, respectively.

3 Experiment 2: obscured focus of expansion

We repeated the first experiment with the constraint that the moving object was on the same side of the screen as the heading point, and obscured the FOE for all or most of every trial. To determine whether simply covering the FOE affects perceived heading, we added a black object that occluded the background dots but possessed no dots itself, and thus introduced no inconsistent dot motion. Thus, if heading is determined by locating or fixating the explicit FOE, performance should be affected similarly by the transparent, opaque, and black objects, whereas if heading is determined by pooling over all dot motions, the transparent and opaque objects should affect performance, but black object should not. In addition, we added a set of screening trials to measure baseline performance without a moving object and to exclude observers who could not perform the basic heading task.

3.1 Method

3.1.1 Observers. Eighteen observers were paid to participate, only one of whom had been in experiment 1. To minimize differences due to inability to perform the basic heading task, observers were screened prior to participating in moving-object trials.

The screening test consisted of a block of 128 trials that presented only the background plane, with probe and heading positions identical to those in experiment 1. Observers were removed from the sample if they had a constant error greater than 2.5° , or if they failed to exceed 75% correct at any heading angle. Six observers were excluded on the basis of these criteria, leaving twelve in the final group. Thus, the results only generalize to individuals who can perform the basic heading task reliably.

3.1.2 Displays. Displays were identical to those in experiment 1, with three exceptions: (a) the initial position of the object was chosen so that it was on the same side of the screen as the probe; (b) only one initial object size was used, with a diameter of 10 deg; (c) in addition to Opaque and Transparent conditions, we included a Black condition in which a black, moving square simply occluded the background dots. The same motion parameters were used in all three conditions, so the displays in the Black condition were identical to those in the Opaque condition, with object dots removed.

This resulted in a two-factor design, object type by path angle. Trials were again blocked by object type and counterbalanced for order, with the same practice procedure. The screening trials and first block of test trials were presented in one session, with the other two blocks presented in a second session. The 128 screening trials were followed by 576 test trials, yielding 64 data points for each of the 9 heading estimates. PSEs and DLs were computed as in experiment 1. 11 of the 108 heading estimates were excluded from the analysis because of poor fits, and 20 estimates with $|PSE| > 8$ were truncated to ± 8 . Because this removed nearly 30% of the data points from the DL analysis, we did not analyze variable error for this experiment.

3.2 Results and discussion

Under these conditions, perceived heading was significantly biased toward the FOE', opposite to the direction of object motion. In figure 4 mean heading error is plotted as a function of path angle for each of the three object types. The opaque and transparent objects produced a heading error that increased with path angle, with a greater effect for the opaque object. On the other hand, the black object did not yield a bias.

A two-way multivariate repeated-measures ANOVA on heading error revealed a main effect of path angle ($F_{2,10} = 12.638, p = 0.002$), no main effect of object type ($F_{2,10} = 0.80, p = 0.461, ns$), but a significant interaction ($F_{4,8} = 12.64, p = 0.002$). The overall trend in heading error as a function of path angle was linear ($F = 26.78, p < 0.001$), with no quadratic component ($F = 0.062, p = 0.807$). Tests of simple effects with a Bonferroni adjustment setting $\alpha = 0.017$ showed a significant linear trend in the Opaque condition ($F_{1,11} = 34.30, p < 0.001$) and the Transparent condition

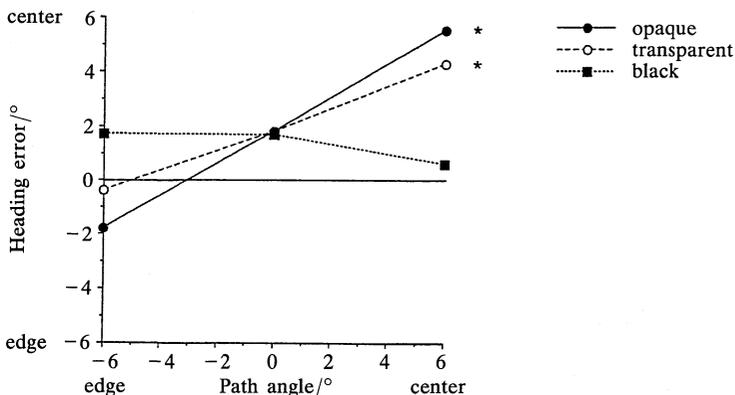


Figure 4. Mean heading error in experiment 2 (obscured FOE) for the three types of object. Asterisks indicate significantly different linear trends from that for the Black condition.

($F_{1,11} = 21.57, p < 0.001$), but not in the Black condition ($F_{1,11} = 5.03, p = 0.047, ns$). The black object thus did not elicit a significant heading bias. Interaction contrasts revealed that the linear trend for the black object was significantly different from that for both the transparent object ($F_{1,11} = 32.53, p < 0.001$) and the opaque object ($F_{1,11} = 44.375, p < 0.001$), although the latter two were not different from each other ($F_{1,11} = 4.952, p = 0.428, ns$). The mean heading error in screening trials with no moving object was 1.54° , showing a slight center-screen bias; in the Black condition it was 1.32° .

In sum, in contrast to in experiment 1, opaque and transparent objects do bias perceived heading in the direction of the FOE' when they obscure the heading point. The mean magnitude of the bias is 2.3° with the transparent object and 3.7° with the opaque object, intermediate between the actual heading point and the full 6 deg displacement of the FOE'. It thus appears that the visual system does not segment moving objects prior to recovering heading, at least on the basis of relative motion or dynamic-occlusion information. Further, the absence of such an effect with the black object indicates that the bias is not due to occlusion of the FOE and the informative vectors around it, but rather to discrepant motion of object dots. These results are consistent with the spatial-pooling hypothesis.

The difference between the results of experiments 1 and 2 suggests that the region of the flow field near the heading point is more susceptible to discrepant motion than regions 6–20 deg away. This is consistent with the notion that vectors near the FOE are more informative and may be more heavily weighted in the heading estimation. Alternatively, it is possible that the results reflect different attentional strategies made possible by blocking the trials. In experiment 1, the object was always on the opposite side of the screen from the heading point, so observers could use it as a cue to the location of the FOE, which could then be fixated quickly. In experiment 2 the FOE was occluded by the object, so observers may have adopted a different strategy of looking above or below the moving object, or attending to the pattern as a whole. In the next experiment, we randomly intermixed the two types of trials.

4 Experiment 3: mixed trials

In experiment 3 we attempted to determine whether the difference in heading bias with a visible and obscured focus of expansion is due to the region of the flow field containing discrepant motion, or to an attentional strategy in which the location of the object was used as a cue to the location of the heading point. We thus randomly intermixed trials in which the object and the heading point were on the same side of the screen (obscured FOE) or on the opposite side of the screen (visible FOE), so the position of the moving object could not be used as a cue.

We also conducted a small control condition to ensure that the displays contained sufficient information to discriminate the moving objects from the background. It is possible that object motion is spatially pooled because it is not discriminable, offering an alternative explanation of the observed heading bias.

4.1 Method

4.1.1 *Observers.* A total of twelve observers were paid to participate, including two from experiment 1 and ten from experiment 2. There was no further screening.

4.1.2 *Displays.* Displays were identical to those of experiment 2, except that the initial position of the object relative to the probe varied randomly from trial to trial. Half the trials were thus in the obscured-FOE condition, in which the object was on the same side of the screen as the heading point, and the other half were in the visible-FOE condition, in which it was on the opposite side of the screen. This resulted in a three-factor design: object type, path angle, and visibility, for a total of 1152 trials

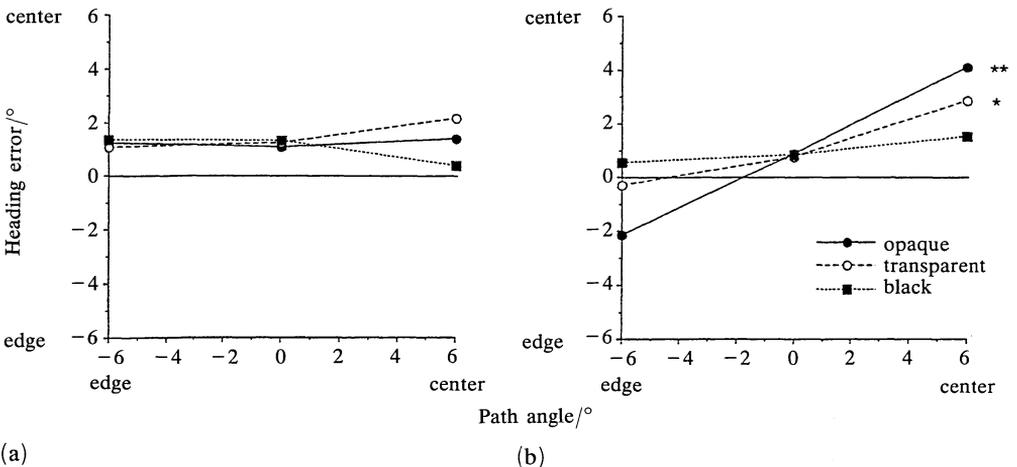
in two sessions. PSEs and DLs were computed as before. 7 of the 216 heading estimates had PSEs < -8 and were truncated to -8; these cases were excluded from the DL analysis.

Two of the observers also participated in an additional control condition, in which they were asked to identify whether or not they saw a coherent moving object. Displays were a subset of those in the Transparent and Opaque conditions, except that the probe line was not presented. There were 192 trials, on half of which there was no moving object.

4.2 Results and discussion

Heading errors for the Visible condition are presented in figure 5a, and for the Obscured condition in figure 5b. The results are similar to those of experiments 1 and 2, respectively: when the heading point is visible, moving objects had no effect (figure 5a), but when it is obscured, both opaque and transparent objects bias perceived heading—this time, the former more so than the latter (figure 5b). An omnibus multivariate repeated-measures ANOVA on heading error revealed that the three-way interaction was significant ($F_{4,8} = 9.07, p < 0.001$), confirming that the interaction between object type and path angle is different in the Visible and Obscured conditions.

We then conducted a separate two-way multivariate ANOVA on heading error in each condition. The Visible condition yielded no significant main effects or interactions, replicating the results of experiment 1. The Obscured condition exhibited a main effect of path angle ($F_{2,10} = 8.92, p = 0.006$), no effect of object type ($F_{2,10} = 0.10, p = 0.90$), but a significant interaction ($F_{4,8} = 6.89, p = 0.011$), replicating the results of experiment 2. Tests of simple effects with a Bonferroni adjustment setting $\alpha = 0.017$ again revealed a significant linear trend in the Opaque condition ($F_{1,11} = 25.25, p < 0.001$) and the Transparent condition ($F_{1,11} = 8.101, p = 0.016$), but not in the Black condition ($F_{1,11} = 2.76, p = 0.125, ns$). Interaction contrasts showed that the linear trend for the black object was significantly different from that for both the opaque object ($F_{1,11} = 26.40, p < 0.001$) and the transparent object ($F_{1,11} = 9.035, p = 0.012$). This time, the trends for the transparent and opaque objects differed as well ($F_{1,11} = 11.080, p = 0.007$), demonstrating that the opaque object produced a significantly greater bias than the transparent object. The mean magnitude of the bias was 1.6° with the transparent object and 3.1° with the opaque object.



(a) (b) **Figure 5.** Mean heading error in experiment 3 for the three types of object: (a) visible FOE; (b) obscured FOE. A single asterisk indicates a significantly different linear trend from that for the Black condition and two asterisks a significant difference from the Black and the Transparent condition.

The overall mean DL was 1.92° , comparable to experiment 1. A three-way repeated-measures ANOVA revealed no significant effects or interactions. This indicates that moving objects induce a predictable bias, not just a greater uncertainty, in the heading direction.

The test of object discriminability revealed that observers could easily identify moving objects in both the Transparent condition (97% correct) and the Opaque condition (100% correct). Thus, the biasing effect of moving objects is not due to the fact that they cannot be discriminated from the background.

In short, the results of present experiment replicate those of experiments 1 and 2. This demonstrates that the difference between the Visible and Obscured conditions is not an artifact of blocking trials, but is likely due to the importance of the region of the flow field around the FOE. In addition, we found that the opaque object induced significantly greater bias than the transparent object. This is again consistent with the spatial-pooling hypothesis, whereas background dots are visible through the transparent object and thus contribute to the heading estimate, the opaque object occluded them, thereby reducing the contribution of the background and increasing the bias.

5 Template model for translational heading

The experimental results appear quite consistent with the spatial-pooling hypothesis. To determine more rigorously whether the principle of central-weighted spatial pooling could account for the data, we implemented a template model based on large-field expansion units, derived from the neural network of Hatsopoulos and Warren (1991).

In the model, sketched in figure 6, we assume an input layer composed of local velocity-selective units, analogous to area MT (Rodman and Albright 1987), and an output layer composed of expansion-selective units with large, center-weighted receptive fields, analogous to cells in area MST (Saito et al 1986). The input layer has a columnar organization, with a set of units sensitive to various directions and speeds represented at each retinal location x . The output layer forms a two-dimensional 'heading map' in retinal coordinates, such that each output unit h responds preferentially to the family of velocity fields $\{v_h(x)\}$ associated with a specific heading direction and a range of depth values (scaled to observer speed). The response $R(h)$ of the unit is weighted by a Gaussian function that emphasizes input vectors near the center of its receptive field. Thus, as illustrated in figure 7, a moving object in front of the FOE will exhibit the activity in unit h_1 corresponding to the actual heading point, and contribute to the activity in a neighboring unit h_2 , thereby shifting the peak of the response distribution toward the FOE'. A natural extension would be to include units

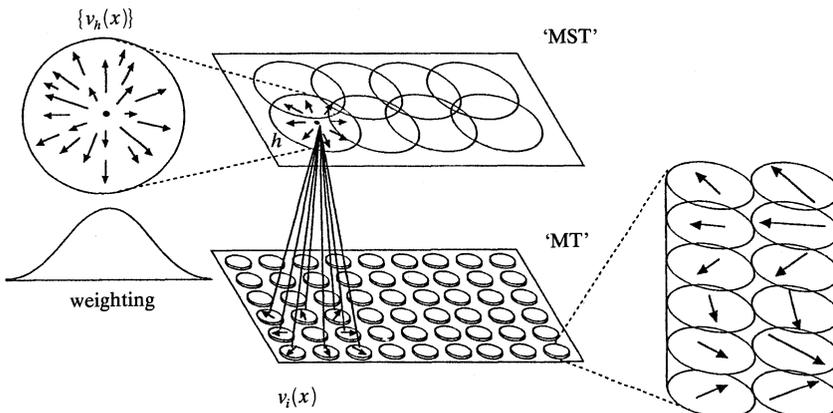


Figure 6. Diagram of the expansion-template model. See text for details.

selective for the family of flow patterns produced by combinations of translation and rotation (Perrone and Stone 1994; Warren, in press).

Center-weighted receptive fields in the output layer were modeled by a Gaussian filter of variance σ_h , centered on location x_h . The width of the Gaussian was constant ($\sigma_h = \sigma$ for all h), and the center of the receptive field was the preferred heading point. The response of these units is a function of how well an input velocity field $v_i(x)$ matches the unit's preferred patterns $\{v_h(x)\}$, weighted by the Gaussian. Since, for observer translation, variations in depth alter the magnitudes but not the directions of flow vectors, the response to input vectors was assumed to be a function of vector direction alone. Each location x in the receptive field of output unit h is associated with a preferred vector direction, and the contribution of an input vector to the response of the unit is the cosine of the angle between the input vector and the preferred vector.⁽¹⁾ For an input velocity field $v_i(x)$, the response of unit h is then:

$$R(h) = \int Z \left(\frac{x - x_h}{\sigma} \right) \frac{v_i \cdot v_h}{|v_i||v_h|} dx, \quad (1)$$

where Z is the standard normal distribution. $R(h)$ is a measure of the overall similarity of the input pattern $v_i(x)$ with a preferred pattern $v_h(x)$ of a given output unit h , with central locations weighted more heavily by the Gaussian filter. Perceived heading can be inferred to be the maximum of this function.

In the following simulations, we used a 40×32 array of input units with a spacing of 1 deg, which sampled the full field of view of our displays. The output layer was a 80×32 array of output units spaced 0.5 deg apart horizontally and 1 deg vertically.

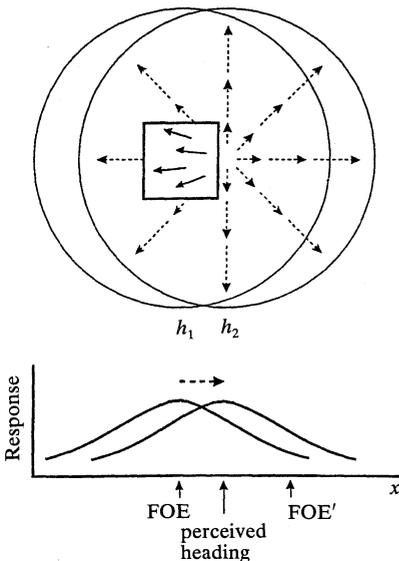


Figure 7. Conceptual illustration of the influence of a moving object in neighboring receptive fields. Unit h_1 is centered on the actual heading point (FOE), the preferred pattern of unit h_2 is indicated by dashed vectors. A moving object at the heading point reduces activity in h_1 and increases activity in h_2 , yielding a shift in the response distribution toward FOE'.

⁽¹⁾ Note that if the model were modified to include observer rotation and translation, the cosine of the difference angle would no longer be an adequate metric, because the contribution of a local velocity vector to the response of h would have to depend on its magnitude as well as its direction.

Output-layer responses were computed directly from equation (1), with the width of the Gaussian filter set at $\sigma = 10$. These parameters were adopted as reasonable values and tested without optimization. For the transparent condition, the superimposed velocity fields of the object and background were assumed to be additive, which ignores possible saturation effects at the input layer. Since we deal only with horizontal shifts in perceived heading, response functions $R(h)$ were summed over all units with preferred heading at the same horizontal position, yielding summed response, $S(x)$, as a function of horizontal position x .

5.1 Performance of the model

We first tested the model with some representative cases from experiment 2. For example, an input velocity field was generated with a heading of 6° for each object in Obscured and Visible conditions. Figure 8 represents the summed response distribution for each condition. With an obscured FOE and a path angle of 0° (figure 8a) the peak response for each object lines up near the heading point, with a slight center-screen bias. With a path angle of 6° (figure 8b), the peak without a moving object is near the FOE, but it shifts toward FOE' by about 1.5 deg in the Transparent condition

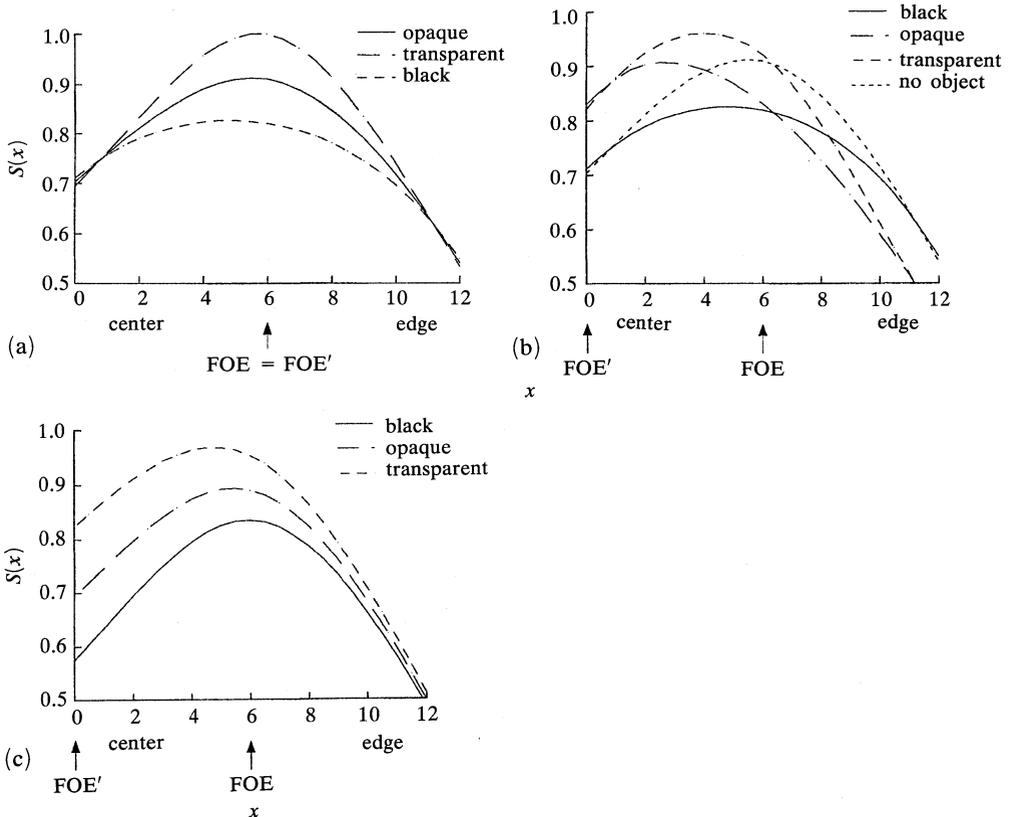


Figure 8. Summed response $S(x)$ as a function of horizontal position x . Responses, $R(h)$, were summed over all output units with preferred headings at the same horizontal position x . The location of the maximum of $S(x)$ can be inferred to be the horizontal position of perceived heading. (a) Obscured FOE, heading = 8° , path angle = 0° . The peak response for each object is close to the actual heading, but shows slight center-screen bias. (b) Obscured FOE, heading = 8° , path angle = 6° . The peak response with no moving object is close to the actual heading (FOE), but those with transparent and opaque objects are biased toward the FOE'. (c) Visible FOE, heading = 8° , center of object = -8° , path angle = 6° . There is little influence of the moving object on the peak response.

tion and 3 deg in the Opaque condition. The shift is greater with the opaque object because object motion increases activity in units toward the FOE', and the occlusion of background motion simultaneously reduces activity in units close to the FOE. The overall response is larger with the transparent object because intermediate output units are activated by both object and background motion. There is also a small shift in the Black condition, which we return to below. Last, with a visible FOE (figure 8c), such that the center of the object is 12 deg from the heading point, all peaks are near the actual heading point, but exhibit a small shift owing to a center-screen bias that is exaggerated by missing or discrepant information.

The model offers an explanation of the frequently observed center-screen bias. If motion information is missing at the edge of the screen, this reduces activity in units near the edge relative to neighboring units toward the center simply because there is less consistent motion in their receptive fields. This results in a small shift in the peak, which gets larger as the heading point approaches the edge of the screen. This effect depends on a broad tuning, such that the receptive fields of many units overlap the edge of the screen. The small shifts seen in the Black condition can be similarly explained. In this case, motion information is missing in the heavily weighted region around the FOE, depressing the overall response and flattening the function. Owing to the heavy weighting of this region, small differences in the position of the missing information can correspondingly shift the peak. The direction and magnitude of the shift thus depends on initial conditions, and yields a small average bias.

We then tested the model on the flow fields of experiment 3. Heading error was taken to be the difference between the actual heading and the location of the peak response, and these errors were determined for all observer-heading and object-motion conditions tested in experiment 3. Figure 9 shows the resulting mean heading error in the Visible and Obscured conditions, which are very similar to the human data from experiment 3 (figure 5). In the Visible condition (figure 9a), path angle has no effect on perceived heading, whereas in the Obscured condition (figure 9b), heading bias increases with path angle. This effect is larger for the opaque object than the transparent object, and absent for the black object, which is consistent with our data. An overall center-screen bias is also evident in both conditions, as expected.

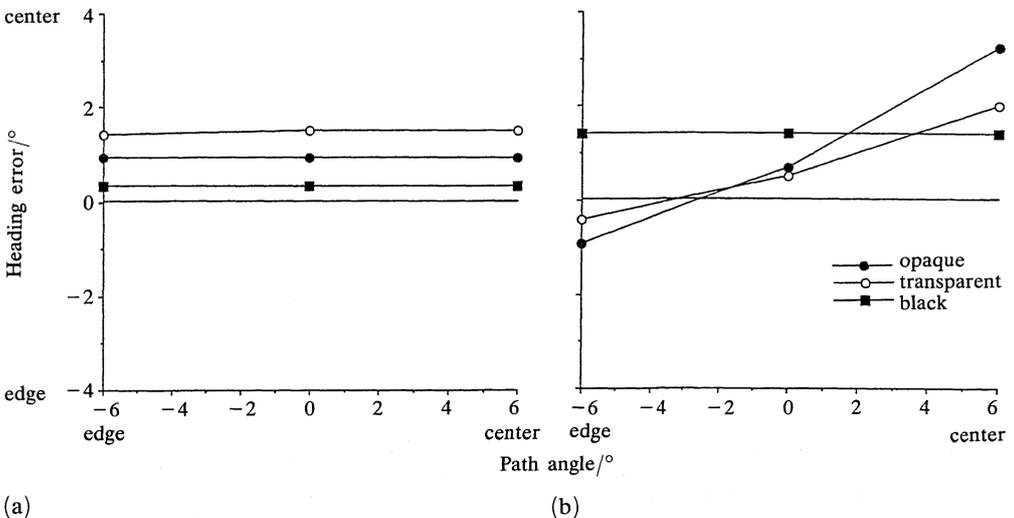


Figure 9. Mean heading error for the model on flow fields from experiment 3. (a) Visible FOE; (b) obscured FOE. Note the increased center-screen bias with the black object owing to missing vectors near the heading point.

But figure 9a shows different amounts of center bias for different object types, which is not present in the human data. The difference between opaque and transparent objects is due to our simplifying assumption of additivity in the input layer for superimposed velocity fields. A more elaborate model that included saturation and threshold responses in the MT layer would reduce this effect. With the black object, the amount of center bias depends on the average position of the missing information in the Visible and Obscured conditions, although this effect does not show up in the human data. However, the key effects of path angle, object type, and FOE visibility are all accounted for by the principle of center-weighted spatial pooling embodied in the model.

6 General discussion

The results show that perceived heading is unaffected by a moving object far from the heading point, but is significantly biased by an object moving in front of the heading point. The bias is toward the FOE' (opposite the direction of object motion) and is greater with an opaque object than a transparent object. The absence of bias with a black object indicates that it is due to the discrepant motion of object dots, not occlusion of the background.

Somewhat surprisingly, this pattern of results indicates that the visual system does not segment moving objects prior to determining heading. Dynamic-occlusion information for segmentation does not aid the heading estimate, for the bias is actually greater with the opaque object. The transparent object also induces significant bias, despite a maximal amount of relative-motion information for the object, contrary to Hildreth's (1992) theory. Segmentation does not fail because the moving object cannot be discriminated, for observers detect the presence of both the transparent and the opaque object with complete accuracy, and the dynamic-occlusion boundary is subjectively quite vivid. It thus appears that segmentation and heading estimation may be functionally separate processes.

These findings are quite consistent with the center-weighted spatial-pooling hypothesis. Perceived heading is intermediate between the FOE and FOE', as predicted if both sets of motion vectors contribute to the heading estimate. This accounts for the greater bias of the opaque object, for eliminating background vectors reduces the influence of the background (FOE) relative to that of the object (FOE'). Conversely, more background vectors are visible through the transparent object, increasing the influence of the background (FOE) relative to the object (FOE') and yielding a smaller bias. Last, the fact that the bias only occurs when objects move in front of the heading point can be accounted for by weighting the most informative vectors near the FOE. Thus, discrepant motion near the FOE will affect the heading estimate more than the same motion far from the FOE.

We implemented this idea in a simple template model based on large-field, center-weighted expansion units. When tested on flow fields from the present experiments, the results closely captured the pattern of the human data. The model also accounts for frequent observations of a center-screen bias in heading judgments. The strength of the model is its simplicity, for it demonstrates that the main features of human performance can be explained by the principle of center-weighted spatial pooling, without other model-dependent assumptions about the details of motion processing.

The model is also consistent with properties of motion-selective cells in area MSTd of the primate visual cortex, as far as they are understood (Saito et al 1986; Tanaka and Saito 1989; Tanaka et al 1989). These units are selective for expansion, rotation, translation, and certain of their combinations, and appear to act as templates for global patterns of motion rather than decomposing the flow pattern into basic components (Duffy and Wurtz 1991a, 1991b; Graziano et al 1994; Orban et al 1992). In particular,

they have large receptive fields, are insensitive to variation in dot density, and do not distinguish local object motion from global field motion, even with clear boundary information (Andersen 1994). Thus, MST units may perform the sort of spatial pooling suggested by our results, although their functional role in motion perception remains uncertain.

Our conclusions must be considered preliminary for several reasons. First, additional surface information (stereo, brightness or color contrast, etc.) might increase the salience of the moving object, yielding segmentation prior to heading. But given that the boundaries of the opaque object are already highly salient, we doubt this would alter the results. Second, heading toward a frontal plane is known to be ambiguous under conditions of simulated eye rotation (Longuet-Higgins 1984; Warren and Hannon 1990), and it has been suggested that the visual system makes use of motion parallax owing to depth variation in the scene (Rieger and Lawton 1985). It is possible that a more complex three-dimensional background might reduce the present bias, and should be tested. Third, to determine the role of eye movements, fixation should be controlled by placing fixation points on the stationary background and the moving object. There are numerous other variables to explore, such as the size of the object, its proximity to the FOE, its direction of motion, and the number of objects.

We should also note that our results appear to conflict with those recently reported by Royden and Hildreth (1994) using a virtually identical paradigm. They too found an effect of a moving object when it covered the FOE, but in the opposite direction: perceived heading was biased in the direction of object motion, not in the opposite direction. We believe this is due to a key difference in their displays, for object motion was predominantly lateral and contained little expansion, whereas in our displays it was predominantly in depth with a large expansion component. Similar 'motion capture' of the FOE was previously reported by Duffy and Wurtz (1993) when they superimposed a translational motion upon an expansion pattern. As observed in MSTd cells, translational motion contributes little to activity in expansion units, but such displays would presumably activate double-component units sensitive to both expansion and translation (Duffy and Wurtz 1991a, 1991b). How this might lead to a perceptual bias in perceived heading is not well understood.

Center-weighted spatial pooling may actually represent a simple, adaptive solution for navigation. Suppose the relevant task for locomotion is not to determine heading in the frame of reference of the stationary surround, but to steer with respect to obstacles on the path ahead. When one drives on the highway, for example, it may be more important to perceive one's heading relative to the vehicle in front (FOE') than the roadway itself (FOE). As the observer approaches a moving object, its visual angle will increase and at some point ($\geq 10^\circ$) may come to dominate perceived heading; but if the object is off the observer's path, it will not. This achieves adaptive control without segmenting objects or recovering their three-dimensional layout and motion parameters. Such a strategy is consistent with the view that steering and perhaps other aspects of navigation may be based on task-specific information, rather than on a general-purpose three-dimensional representation of the scene (Aloimonos 1993; Brooks 1991; Warren 1988).

Acknowledgments. This research was supported by grant AG05223 from the National Institutes of Health. We would like to thank Richard Hlustick for programming the original displays.

References

- Adiv G, 1985 "Determining three-dimensional motion and structure from optical flow generated by several moving objects" *IEEE Pattern Analysis and Machine Intelligence* 7 384-401
- Aloimonos Y, 1993 *Active Perception* (Hillsdale, NJ: Lawrence Erlbaum Associates)
- Andersen R A, 1994, paper presented at the Workshop on Systems-Level Models of Visual Behavior, July, Telluride, CO

- Berg A V van den, 1992 "Robustness of perception of heading from optic flow" *Vision Research* **32** 1285–1296
- Brooks R A, 1991 *Intelligence without Reason* AI memo no. 1293, MIT Artificial Intelligence Laboratory, Boston, MA
- Bruss A R, Horn B K P, 1983 "Passive navigation" *Computer Vision, Graphics and Image Processing* **21** 3–20
- Crowell J A, Banks M S, 1993 "Perceiving heading with different retinal regions and types of optic flow" *Perception & Psychophysics* **53** 325–337
- da Vitoria Lobo N, Tsotsos J K, 1991 "Telling where one is heading and where things move independently", in *Proceedings of the 13th Annual Conference of the Cognitive Science Society, Chicago, IL* (Hillsdale, NJ: Lawrence Erlbaum Associates) pp 43–48
- Duffy C J, Wurtz R H, 1991a "Sensitivity of MST neurons to optic flow stimuli. I. A continuum of response selectivity to large-field stimuli" *Journal of Neurophysiology* **65** 1329–1345
- Duffy C J, Wurtz R H, 1991b "Sensitivity of MST neurons to optic flow stimuli. II. Mechanisms of response selectivity revealed by small-field stimuli" *Journal of Neurophysiology* **65** 1346–1359
- Duffy C J, Wurtz R H, 1993 "An illusory transformation of optic flow fields" *Vision Research* **33** 1481–1490
- Gibson J J, 1950 *Perception of the Visual World* (Boston, MA: Houghton Mifflin)
- Graziano M S A, Andersen R A, Snowden R J, 1994 "Tuning of MST neurons to spiral motions" *Journal of Neuroscience* **14** 54–67
- Hatsopoulos N G, Warren W H, 1991 "Visual navigation with a neural network" *Neural Networks* **4** 303–317
- Heeger D J, Hager G, 1988 "Egomotion and the stabilized world", in *Proceedings of the 2nd International Conference on Computer Vision, Tampa, FL* (Washington, DC: IEEE) pp 435–440
- Heeger D J, Jepson A D, 1992 "Subspace methods for recovering rigid motion I: Algorithm and implementation" *International Journal of Computer Vision* **7** 95–117
- Hildreth E, 1992 "Recovering heading for visually-guided navigation" *Vision Research* **32** 1177–1192
- Koenderink J J, Doorn A J van, 1987 "Facts on optic flow" *Biological Cybernetics* **56** 247–254
- Lappe M, Rauschecker J P, 1993 "A neural network for the processing of optic flow from egomotion in man and higher mammals" *Neural Computation* **5** 374–391
- Longuet-Higgins H C, 1984 "The visual ambiguity of a moving plane" *Proceedings of the Royal Society of London, Series B* **223** 165–175
- Longuet-Higgins H C, Prazdny K, 1980 "The interpretation of a moving retinal image" *Proceedings of the Royal Society of London, Series B* **208** 385–397
- Orban G A, Lagae L, Verri A, Raiguel D, Xiao D, Maes H, Torre V, 1992 "First-order analysis of optical flow in monkey brain" *Proceedings of the National Academy of Sciences of the United States of America* **89** 2595–2599
- Perrone J A, 1992 "Model for the computation of self-motion in biological systems" *Journal of the Optical Society of America A* **9** 177–194
- Perrone J A, Stone L S, 1994 "A model of self-motion estimation within primate extrastriate visual cortex" *Vision Research* **34** 2917–2938
- Rieger J H, Lawton D T, 1985 "Processing differential image motion" *Journal of the Optical Society of America A* **2** 354–360
- Rodman H R, Albright T D, 1987 "Coding of visual stimulus velocity in area MT of the Macaque" *Vision Research* **27** 2035–2048
- Royden C S, Hildreth E C, 1994 "The effects of moving objects on heading perception" *Investigative Ophthalmology and Visual Science, Supplement* **34** 1999
- Royden C S, Banks M S, Crowell J A, 1992 "The perception of heading during eye movements" *Nature (London)* **360** 583–585
- Royden C S, Crowell J A, Banks M S, 1994 "Estimating heading during eye movements" *Vision Research* **34** 3197–3214
- Saito H, Yuki M, Tanaka K, Hikosaka K, Fukada Y, Iwai E, 1986 "Integration of direction signals of image motion in the superior temporal sulcus of the Macaque monkey" *Journal of Neuroscience* **6** 145–157
- Tanaka K, Saito H, 1989 "Analysis of motion of the visual field by direction, expansion/contraction, and rotation cells clustered in the dorsal part of the medial superior temporal area of the Macaque monkey" *Journal of Neurophysiology* **62** 626–641
- Tanaka K, Fukada Y, Saito H, 1989 "Underlying mechanisms of the response specificity of expansion/contraction and rotation cells in the dorsal part of the medial superior temporal area of the Macaque monkey" *Journal of Neurophysiology* **62** 642–656

-
- Thompson W, Pong T, 1990 "Detecting moving objects" *International Journal of Computer Vision* **4** 39-57
- Thompson W B, Lechleider P, Stuch E R, 1993 "Detecting moving objects using the rigidity constraint" *IEEE Transactions on Pattern Analysis and Machine Intelligence* **15** 162-165
- Tsai R Y, Huang T S, 1981 "Estimating three-dimensional motion parameters of a rigid planar patch" *IEEE Transactions on Acoustics, Speech and Signal Processing* **29** 1147-1152
- Warren W H, 1988 "Action modes and laws of control for the visual guidance of action", in *Movement Behavior: The Motor-Action Controversy* (Amsterdam: North-Holland) pp 339-380
- Warren W H, in press "Self-motion: Visual perception and visual control", in *Handbook of Perception and Cognition* volume 5: *Perception of Space and Motion* Eds W Epstein, S Rogers (San Diego, CA: Academic Press)
- 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, Kurtz 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, Blackwell A W, Kurtz K J, Hatsopoulos N G, Kalish M L, 1991 "On the sufficiency of the velocity field for perception of heading" *Biological Cybernetics* **65** 311-320
- Warren W H, Morris M W, Kalish M, 1988 "Perception of translational heading from optical flow" *Journal of Experimental Psychology: Human Perception and Performance* **14** 646-660
- Waxman A M, Ullman S, 1985 "Surface structure and 3D motion from image flow: a kinematic analysis" *International Journal of Robotics Research* **4** 72-94
- Zhang Z, Faugeras O D, Ayache N, 1988 "Analysis of a sequence of stereo scenes containing multiple moving objects using rigidity constraints", in *Proceedings of the 2nd International Conference on Computer Vision, Tampa, FL* (Washington, DC: IEEE) pp 177-186