Two-Dimensional Motion Perception in Flies

A. Borst
M. Egelhaaf
(Max-Planck-Institut für biologische Kybernetik,
Spemannstrasse 38, 7400 Tübingen, Germany)

H. S. Seung*
(Racah Institute of Physics and Center for Neural Computation,
Hebrew University, Jerusalem 91904, Israel)

We study two-dimensional motion perception in flies using a semicircular visual stimulus. Measurements of both the H1-neuron and the optomotor response are consistent with a simple model supposing spatial integration of the outputs of correlation-type motion detectors. In both experiment and model, there is substantial H1 and horizontal (yaw) optomotor response to purely vertical motion of the stimulus. We conclude that the fly's optomotor response to a two-dimensional pattern, depending on its structure, may deviate considerably from the direction of pattern motion.

1 Introduction

The projection of the velocity vectors of objects moving in three-dimensional space on the image plane of an eye or a camera can be described as a vector field. This two-dimensional velocity field is time-dependent and assigns the direction and magnitude of a velocity vector to each point in the image plane. However, the velocity field is a purely geometric concept and does not directly represent the input of a visual system. Instead, the only information available to a visual system is given by the time-dependent brightness values as sensed by photoreceptors in the image plane. The problem of motion perception has often been posed as that of recovering the velocity field from these brightness values.

For the case of simple translation of a Lambertian surface under uniform illumination, this computation can be done (Verri and Poggio 1989). Such a physical motion leads to the translation of a brightness pattern across the image plane. Several known local algorithms (Reichardt et al. 1988; Srinivasan 1990; Uras et al. 1988) recover the correct velocity field, which is constant in space and time. Algorithms utilizing a smoothness

*Present address: AT&T Bell Laboratories, Murray Hill, NJ 07974 USA.
regularizer (Horn and Schunck 1981; Hildreth and Koch 1987) also perform well in extracting the true velocity. All these algorithms fail to yield a unique answer only for the special case of one-dimensional patterns. This is because a moving one-dimensional pattern is consistent with an infinite number of velocity fields (Ullman 1983; Horn and Schunk 1981; Hildreth and Koch 1987; Reichardt et al. 1988).

In contrast, the superposition of two differently oriented one-dimensional sine gratings, a plaid pattern, has a uniquely determined velocity vector. The direction of motion of such a two-dimensional (2d) pattern, which is different from the orientations of its one-dimensional (1d) component gratings, is perceived by human observers under certain conditions (Adelsen and Movshon 1982; Ferrera and Wilson 1990; Stone et al. 1990). On the basis of physiological experiments with plaid stimuli, Movshon and co-workers (1986) have argued that motion processing in the primate visual system takes place in two stages. The first stage is composed of local movement detectors in areas V1 and MT sensitive to the orientation of the components. The second stage of processing is composed of neurons in MT that respond to the direction of pattern motion, presumably computing it from the output of the first stage.

In this work, we examine whether the fly visual system computes the direction of motion of a 2d pattern. In the past, the fly has proven to be an excellent model system for analyzing motion detection. Most notably, much is known about the structure and physiology of its visual ganglia and the motion-dependent behaviors controlled by them (for review, see Egelhaaf et al. 1988; Hansen and Egelhaaf 1989; Egelhaaf and Borst 1993). One such motion-dependent behavior is the optomotor response, in which the fly tends to move so as to stabilize a moving visual surround (Fermi and Reichardt 1963; Götz 1972). A simple model of the fly's optomotor pathway has been quite successful in accounting for both neurophysiological and behavioral data. According to this integrated correlation model, there are local movement detectors of the correlation type (Fig. 1A) (Hassenstein and Reichardt 1956; Reichardt 1961; Borst and Egelhaaf 1989) that are organized in two-dimensional retinotopic arrays covering the entire visual field. A set of identified, directionally selective, large-field interneurons in the third visual ganglion spatially integrates over the responses of horizontally oriented detectors in this array (Hansen and Egelhaaf 1989). The yaw optomotor response is a low-pass filtered version of the output of this horizontal system (Egelhaaf 1987). There is also a vertical system in the third visual ganglion, believed to mediate the pitch optomotor response according to an analogous model (Hengstenberg 1982).

The visual pattern used in our experiments was a dark circular disk moving on a bright background. The predictions of the integrated correlation model were compared with the responses of the H1-neuron (a cell
Figure 1: Outline of the motion detector model used to derive the predictions shown in Figures 2 and 3. (A) Single correlation-type motion detector consisting of two mirror-symmetrical subunits with opposite preferred directions. Each subunit has two input lines. The signal in one line is temporally filtered and then multiplied with the unfiltered signal of the other line. The outputs of the subunits are subtracted from each other. (B) Responses of a two-dimensional array of orthogonally oriented pairs of motion detectors to a disk moving in the $y$ direction. Shown is the vector field of equation 2.3, which was calculated using the continuum approximation of Reichardt (1987). The components of each vector are the responses of the $x$- and $y$-detectors at that point. The response is only nonzero on the boundary of the disk.
integrating horizontally oriented motion detectors) and the optomotor response about the vertical axis (yaw torque).

2 Responses of Correlation-Type Motion Detectors

Our model consists of a two-dimensional square lattice of correlation-type motion detectors. At each point of the lattice is a pair of detectors that is oriented along the x and y axes. The luminance of the stimulus is denoted by \( I(r, t) \), where \( r \) is a vector in the \( xy \) plane and \( t \) is time. We treat the responses of the detector pair at \( r \) as the components of a vector, \( \mathbf{d}(r, t) \)

\[
\mathbf{d}(r, t) = \begin{bmatrix}
I(r, t)I(r + e_x, t + \Delta t) - I(r, t + e_x, t)I(r + \Delta t, t) \\
I(r, t)I(r + e_y, t + \Delta t) - I(r + e_y, t)I(r, t + \Delta t)
\end{bmatrix}
\] (2.1)

Here \( e_x \) and \( e_y \) are vectors in the x and y directions, corresponding to the spacing between adjacent lattice points. The two terms of each component of \( \mathbf{d} \) correspond to the opponent subunits shown in Figure 1A, each a spatiotemporal cross-correlation of luminances. In equation 2.1 the temporal filtering is written as a simple time delay \( \Delta t \), but in our computer simulations was more realistically modeled as a low-pass filter. The response of the H1-neuron is modeled as the x component of the integrated response vector,

\[
\mathbf{D}(t) = \int dr \mathbf{d}(r, t)
\] (2.2)

The yaw torque of the fly is modeled as a low-pass filtered version of the neural response. We chose a time constant of 3 sec, which is consistent with Egelhaaf (1987) and yields a good fit to the experimental data.

Figure 1B shows that the local response \( \mathbf{d}(r, t) \) to a upwardly moving circular disk is very different from its velocity field. The velocity field (not shown) is everywhere constant and points in the y direction. The local response, on the other hand, varies greatly: it is zero inside and outside the disk, and takes the form

\[
\mathbf{d}(\theta) \propto \hat{r} \sin \theta
\] (2.3)

on the boundary. Here \( \theta \) denotes the angle from the x axis, and \( \hat{r} \) is the unit vector in the radial direction. This formula follows from equation 4.4 of the appendix, where the local response is calculated analytically using a continuum approximation (Reichardt and Guo 1986; Reichardt 1987; Reichardt and Schlögl 1988). It is evident from Figure 1B that the x-detectors bias the direction away from the true velocity. At the upper right and lower left edges they signal positive (rightward) motion whereas at the opposite sides they signal negative (leftward) motion. Thus the circle in vertical motion mimics horizontal motion at its obliquely oriented edges.

\(^1\)A more realistic triangular lattice model of the fly photoreceptor array (Buchner 1976; Buchner et al. 1978) yields similar predictions.
This is not surprising, since for this stimulus the brightness in both input channels of a horizontal motion detector is either increased or decreased during vertical motion in the same temporal sequence as during horizontal motion.

The pattern dependence of the local response is manifest, and has been studied previously (Reichardt 1987; Borst and Egelhaaf 1989). However, the flight control system of the fly is thought to depend on the integrated response of such an array of motion detectors (Egelhaaf et al. 1989). Although there is significant local x-response to vertical motion of a circle (Fig. 1B), the integrated x-response is exactly zero because the contributions from the left and right halves of the circle cancel each other out. Hence, for a full circular stimulus the direction of the integrated response of an array of correlation-type motion detectors is the same as that of a true velocity sensor.

To create a stimulus without such cancellation effects, the circle was moved behind a square aperture in such a way that maximally only a semicircle was visible. Figures 2 and 3A exhibit the analytic results and numerical simulations for this semicircular stimulus. Two features of the model predictions are noteworthy: (1) In contrast to a true velocity sensor, the integrated x-output responds not only to horizontal motion but also to vertical motion, with a time course that depends on the stimulus pattern. (2) The response to horizontal motion remains the same sign throughout the duration of the stimulus, the sign depending on the direction of motion. In contrast, the response to vertical motion changes sign when half of the semicircle is visible in the aperture, thereby erroneously mimicking an inversion of the direction of motion.

### 3 Responses of the Fly

These predictions were first tested by recording the spike activity of the H1-neuron in female blowflies (Calliphora erythrocephala) following standard procedures (Hausen 1982; Borst and Egelhaaf 1990). The resulting spike-frequency histograms of response to a moving semicircular stimulus are shown in Figure 3B. The preferred direction of the H1-neuron, when tested with 1d grating patterns, is horizontal motion from back to front in the entire visual field of one eye (McCann and Dill 1969; Hansen 1976). Similarly, we found that the neuron was excited by horizontal motion of the semicircular stimulus in the preferred direction, and slightly inhibited by motion in the null direction. The magnitude of the null response is smaller than that of the preferred response, probably due to the low spontaneous activity of the cell and the resulting rectification nonlinearity. The response of the H1-neuron to vertical motion of a 1d grating pattern is negligible (Hansen 1976). However, during vertical motion of the semicircular stimulus, the neuron's response shows pronounced modulations and even a sign reversal relative to the resting level (hori-
Figure 2: Spatially integrated response of correlation-type motion detectors to a moving circle seen through an aperture, calculated in equations A.8 and A.9 using the continuum approximation of Reichardt (1987). The $x$-response to motion in the $y$-direction (transverse response), takes the form $M_{xy}(\gamma) = \alpha \left[ 1 - (1 - 2\gamma)^2 \right]$, and the $x$-response to motion in the $x$-direction (longitudinal response) takes the form $M_{xx}(\gamma) = \alpha \cos^{-1}(1 - 2\gamma)$. The prefactor $\alpha$ has a quadratic dependence on the contrast of the stimulus, and $\gamma$ is the fraction of the semicircle that is visible in the aperture. These formulas are valid for the first half of the stimulus period, when $\gamma$ is increasing from 0 to 1. The formulas for the second half of the period are similar. The stimulus trace indicates the visible part of the circle at various instances of time.

Horizontal line in Figure 3B). The responses to upward and downward motion are modulated in the same way but have opposite signs. Except for the fact that the neuron’s responses to vertical motion are almost as strong as its responses to horizontal motion, all response features are in good agreement with the predictions of the correlation-detector model (compare Fig. 3B with Figs. 3A and 2). Hence, the response of the H1-neuron
Figure 3: (A) Spatially integrated responses of a square lattice of correlation-type motion detectors to a black circle moving in various directions on a bright background behind a square aperture. A 20×20 array of motion detectors of the correlation type (Fig. 1) was simulated on an IBM PS/2 using the ASYST language (Keithley Instruments). The motion detectors had a sampling base of one lattice constant, a first-order low-pass filter as delay line, and were horizontally oriented with preferred direction from left to right. Note that the responses to horizontal and vertical motion are scaled differently. (B) Responses of an identified, directionally selective, motion-sensitive, large-field neuron (H1-cell) of the blowfly *Calliphora erythrocephala* to the same stimulus. Stimuli were generated on a CRT (Tektronix 608) controlled by an IBM AT through an image synthesizer (Picasso, Innisfree Inc.) with 200 Hz frame rate. The luminance of the circle was 4 cd/m², and that of the background was 24 cd/m². The contrast \((I_{\text{max}} - I_{\text{min}})/(I_{\text{max}} + I_{\text{min}})\) amounted to 71%. The square aperture had the same extent as the diameter of the circle. The stimulus was presented to only the left eye of the fly at a distance of 7 cm. The circle had a diameter of 70° as seen by the fly. The center of the aperture was at 35° horizontal position and 0° vertical position with respect to the fly. Shown are the mean spike frequency histograms (40 ms binwidth) ± the SEM of the recordings of the H1 responses of 10 flies. Each fly was tested between 50 and 100 times (71 times on average) for each stimulus condition. The cell had rightward motion as its preferred direction. The horizontal line marks the resting firing level. The stimulus trace indicates the visible part of the circle at various instants in time.

is not simply the horizontal component of pattern motion. Measuring from neurons in the vertical system would presumably produce analogous results. We can conclude that the large-field cells in the third visual
ganglion of the fly do not represent the $x$ and $y$ coordinates of the pattern motion vector.\textsuperscript{2}

This finding, however, does not rule out the possibility that the $x$ and $y$ components of pattern motion are computed at some later processing stage in the motion pathway of the fly. Therefore, we recorded the fly's behavioral turning responses about its vertical axis.\textsuperscript{3} These measurements were done on female flies of the species \textit{Musca domestica} suspended from a torque-meter (Gotz 1964) following standard procedures (Egelhaaf 1987). The signals of the computer simulations shown in Figure 3A were passed through a first-order low-pass filter with a time-constant of 3 sec to account for the experimentally established low-pass filter between the third visual ganglion and the final motor output (Egelhaaf 1987). This leads to smoothing and phase shifting of the original signal (compare Fig. 4A with 3A). As was found for the H1-neuron, the behavioral responses induced by the semicircle moving either horizontally or vertically are almost perfectly mimicked by the computer simulations (Fig. 4). Again, pronounced responses are induced not only during horizontal motion but also during vertical motion. The latter responses show a quasisinusoidal modulation and, hence, the same sign reversal observed in the H1-response and in the simulations.

4 Conclusions

In principle it is possible to compute pattern velocity from the output of an array of correlation-type motion detectors (Reichardt et al. 1988; Reichardt and Schlögl 1988), provided that the second spatial derivatives of the pattern are also available and nonzero. Nevertheless, we find no evidence of such a computation in the fly; the output of its local motion detectors appears to undergo no more than a simple spatial integration and temporal filtering. Consequently, depending on the structure of the stimulus pattern, the direction of the optomotor response is not generally the same as the direction of pattern velocity.

Since the function of the optomotor response is believed to be course stabilization, it might seem a deficiency for the response to be in the "wrong" direction. How can an organism such as the fly that is able to perform fast and virtuosic visually guided flight maneuvers afford to confound different directions in such a dramatic way?

\textsuperscript{2}The response of the H1-neuron to the vertical motion of a full circle was also measured. Contrary to the predictions of our simple model, there was some small nonzero response. Refinements of the model can be introduced to account for such incomplete cancellation of response, such as unbalanced subtraction of the two subunits (Egelhaaf et al. 1989) and/or spatially inhomogeneous sensitivity (Hansen 1982).

\textsuperscript{3}Unlike the H1 experiments, two copies of the stimulus were presented simultaneously, one to each eye of the fly. Since the optomotor response integrates signals from both eyes, the flicker response is thereby cancelled, leaving only the motion-selective response, which is of interest here. Duplication of the stimulus would have been irrelevant in the H1 experiments, since H1 receives almost exclusively monocular input.
Figure 4: (A) The integrated responses shown in Figure 3A of a two-dimensional array of correlation-type motion detectors, but fed through a first-order low-pass filter with a 3 sec time constant. Note that responses to horizontal and vertical motion are differently scaled. (B) Averaged optomotor turning responses (± SEM) obtained from 10 flies of the species Musca domestica each tested 20 times for each stimulus condition. Clockwise turning tendencies are shown as positive signals, and counterclockwise as negative signals. The stimulus trace indicates the visible part of the circle at various instants in time. Stimulus conditions were the same as for the electrophysiological recordings (Figure 3B) except for the following: (1) Stimuli were presented on either side of the fly. (2) The square aperture had an extent of 60° as seen by the fly. (3) The aperture was centered at 45° horizontal position and 0° vertical position with respect to the fly.

A plausible answer is that under natural conditions, the fly does not confound directions as dramatically as it does with our artificial stimulus. For the great majority of ecologically relevant stimuli, it may be that the spatially integrated response is very close to the direction of motion. Recall that the symmetry of the full circle led to exact cancellation of the simulated transverse response. For a natural pattern, such exact cancellation is no doubt rare, but there may be an approximate cancellation due to statistical averaging over the complex shapes in the pattern.

Appendix: Continuum Approximation

Consider an image that consists of a circle of radius \( r_0 \) with luminance 1 surrounded by a background of luminance 0. This can be written in polar coordinates as

\[
I(r, \theta) = \Theta(r_0 - r)
\]  
(A.1)
where $\Theta(x)$ is the Heaviside step function. Because of the spatial low-pass properties of the fly eye, the effective input to the detector array is a smoothed form of (A.1), which we can write as

$$I(r, \theta) = f(r)$$

(A.2)

The precise form of $f$ is not important in what follows. What is important is that the radius $r_0$ is much larger than the scale of the smoothing, so that $f'(r)$ is negligible except for $r \approx r_0$.

This visual stimulus, moving at velocity $v$, is input to an array of orthogonally oriented detector pairs. The response of a pair is given by equation 2.1. Each detector has sampling base $\Delta x$ and delay time $\Delta t$. In the continuum approximation to equation 2.1, the local response $d(x, y)$ is related to the velocity vector $v$ by an expression of the form (Reichardt 1987; Reichardt and Schloegl 1988)

$$\begin{pmatrix} d_x \\ d_y \end{pmatrix} = \Delta x \Delta t \begin{pmatrix} m_{xx} & m_{xy} \\ m_{yx} & m_{yy} \end{pmatrix} \begin{pmatrix} v_x \\ v_y \end{pmatrix}$$

(A.3)

For a circular stimulus, the response matrix $m$ is

$$m(r, \theta) = \begin{pmatrix} \cos^2 \theta & \sin \theta \cos \theta \\ \sin \theta \cos \theta & \sin^2 \theta \end{pmatrix} a(r) + \begin{pmatrix} -\sin^2 \theta & \sin \theta \cos \theta \\ \sin \theta \cos \theta & -\cos^2 \theta \end{pmatrix} b(r)$$

(A.4)

where $a(r) = f'(r)^2 - f(r)f''(r)$, and $b(r) = f(r)f'(r)/r$. The off-diagonal element $m_{xy}$, the transverse response, is of special interest. It is the response of the $x$-detector to motion in the $y$-direction. The diagonal element $m_{xx}$ is the longitudinal response, i.e., the response of the $x$-detector to motion in the $x$-direction.

Assuming that the detector array is a square lattice of spacing $\Delta x$, the integrated output is

$$D = \frac{\Delta t}{\Delta x} Mv \equiv \frac{\Delta t}{\Delta x} \left( \int \mathbf{m}(r, \theta) r \, dr \, d\theta \right) v$$

(A.5)

where $M$ is the integrated response over the portion of the circle that is visible.

If the full circle is visible, the off-diagonal elements of the integrated response vanish, so that

$$M = \alpha \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

(A.6)

where

$$\alpha \approx \pi \int_0^\infty dr \, r a(r)$$

(A.7)
The integral of the \( b(r) \) term in A.4 has been neglected, since it is much smaller than the \( a(r) \) term. Because \( M \) is proportional to the identity matrix, the integrated response vector \( D \) is in the same direction as the stimulus velocity \( v \).

For the stimulus used in these experiments, a semicircle moving behind a square aperture (shown in Fig. 2), the integrated response matrix is

\[
M_{xy}(\gamma) = \alpha \left[ 1 - (1 - 2\gamma)^2 \right] \\
M_{x\pi}(\gamma) = \alpha \cos^{-1}(1 - 2\gamma)
\]  

(A.8)  

(A.9)

where \( \gamma \propto t \) is the fraction of the semicircle that is visible in the aperture. These formulas are valid for the first half of the stimulus period, when \( \gamma \) is increasing from 0 to 1. The formulas for the second half of the period are derived similarly. The full response curves are shown in Figure 2.

Acknowledgments

We are grateful to K. G. Götz, W. Reichardt, and J. M. Zanker for carefully reading the manuscript. We also thank the people from the summer 1990 Woods Hole course "Neural Systems and Behavior," where this work was started, for the supportive and stimulating atmosphere. We especially thank B. Mensh and C. Gilbert for assistance in the early stages of this investigation.

References


Received 28 August 1992; accepted 10 March 1993.
This article has been cited by: