

This is the web version of a paper that appeared in *Biological Cybernetics*, Vol. 93, No. 4, pp. 275-287, 2005. The original publication is available at www.springerlink.com:

<http://www.springerlink.com/openurl.asp?genre=article&eissn=1432-0770&volume=93&issue=4&spage=275>

Velocity constancy and models for wide-field visual motion detection in insects

Patrick A. Shoemaker¹, David C. O'Carroll², Andrew D. Straw^{2,3}

¹Tanner Research, Inc., 2650 East Foothill Blvd., Pasadena CA 91107, USA

²Discipline of Physiology, University of Adelaide, Adelaide SA 5005, Australia

³Present Address: California Institute of Technology, Mailcode 138-78, Pasadena CA 91125, USA

Abstract. The tangential neurons in the lobula plate region of the flies are known to respond to visual motion across broad receptive fields in visual space. When intracellular recordings are made from tangential neurons while the intact animal is stimulated visually with moving natural imagery, we find that neural response depends upon speed of motion but is nearly invariant with respect to variations in natural scenery. We refer to this invariance as *velocity constancy*. It is remarkable because natural scenes, in spite of similarities in spatial structure, vary considerably in contrast, and contrast dependence is a feature of neurons in the early visual pathway as well as of most models for the elementary operations of visual motion detection. Thus, we expect that operations must be present in the processing pathway that reduce contrast dependence in order to approximate velocity constancy. We consider models for such operations, including spatial filtering, motion adaptation, saturating nonlinearities, and nonlinear spatial integration by the tangential neurons themselves, and evaluate their effects in simulations of a tangential neuron and precursor processing in response to animated natural imagery. We conclude that all such features reduce interscene variance in response, but that the model system does not approach velocity constancy as closely as the biological tangential cell.

1 Introduction

The tangential neurons in the lobula plate of the dipterans (true flies) are among the most widely studied sensory interneurons in biology. These cells are sensitive to visual motion across broad swaths of the visual field, and almost certainly provide information about the state of self-motion of the organism. Some, for example, have been linked to stabilization about the yaw axis in hovering flight in certain species of flies (Hausen and Egelhaaf, 1989). They respond to moving stimuli by integrating the outputs of large numbers of local elementary motion detectors (EMDs), which are retinotopically distributed and situated at an earlier stage of the visual pathway (Krapp *et al.*, 1998; Egelhaaf *et al.*, 1989; Franceschini *et al.*, 1989). Because the tangential cells are large and amenable to intracellular electrophysiological recording, they have been the source of much indirect physiological evidence regarding the EMD. Although some recent work has attempted to identify and model the neuronal basis of the EMD itself (Higgins *et al.*,

2004), this has proved hard to pinpoint due to the technical difficulty of recording from the very small neurons that are believed to be involved.

Prominent among the tangential cells are the identified neurons that comprise the 'horizontal system' (HS) and 'vertical system' (VS) found in a number of species of flies (Hausen, 1993; Hausen 1982). As their names suggest, these neurons are sensitive primarily to horizontal and vertical optic flow, respectively. They respond with graded membrane potentials to such stimuli and show both depolarizing and hyperpolarizing responses, depending on the direction of motion. The direction of optic flow that leads to the greatest depolarization is called the preferred direction; motion in the opposite direction typically elicits hyperpolarization and is called null- or antipreferred-direction motion. Evidence suggests, however, that these cells are not tuned for uniform unidirectional optic flow; their sensitivity to local motion cues has in some cases been mapped out, and both the direction and magnitude of maximal sensitivity show certain variations across the receptive field (Krapp *et al.*, 1998). The resemblance of some of these maps to the patterns of optic flow induced by particular modes of self-motion has led to the hypothesis that the corresponding HS and VS neurons may act as *matched filters* for these patterns (Krapp *et al.*, 1998; Franz and Krapp, 1998), and thus possibly as indicators of particular states of self-motion.

In spite of extensive study of these neurons and of the photoreceptors and other neurons earlier in the visual pathway, new aspects of their physiology continue to come to light, and many details of the processing that they perform are not completely understood. In the laboratory of one of the authors (D.O'C.), we have studied by intracellular recording the responses of HS and VS cells in intact flies that are stimulated visually with moving imagery obtained from natural scenery. We find that, for a particular mode of motion (e.g., simulated yaw rotation), the responses of these cells depend on the speed of motion, but are remarkably invariant with respect to differing natural scenes, even when those scenes vary significantly in contrast and in the structure of objects within them (Straw *et al.*, 2005). This is illustrated below in Fig. 1, which depicts velocity tuning curves obtained with different animated natural scenes from tangential cells in the hoverfly *Eristalis tenax*. We refer to this property as *velocity constancy*, in analogy with color constancy in the human visual system. It appears that HS and VS neurons encode information about the particular pattern of optic flow and its absolute scaling (i.e., overall speed), but that they are largely capable of rejecting the effects of differences in luminance, contrast, and spatial structure in natural scenes. This is remarkable because these parameters are fundamental to early visual processing, and in addition, models for the EMD tend to depend on the spatial structure of a moving stimulus as well as superlinearly on its contrast.

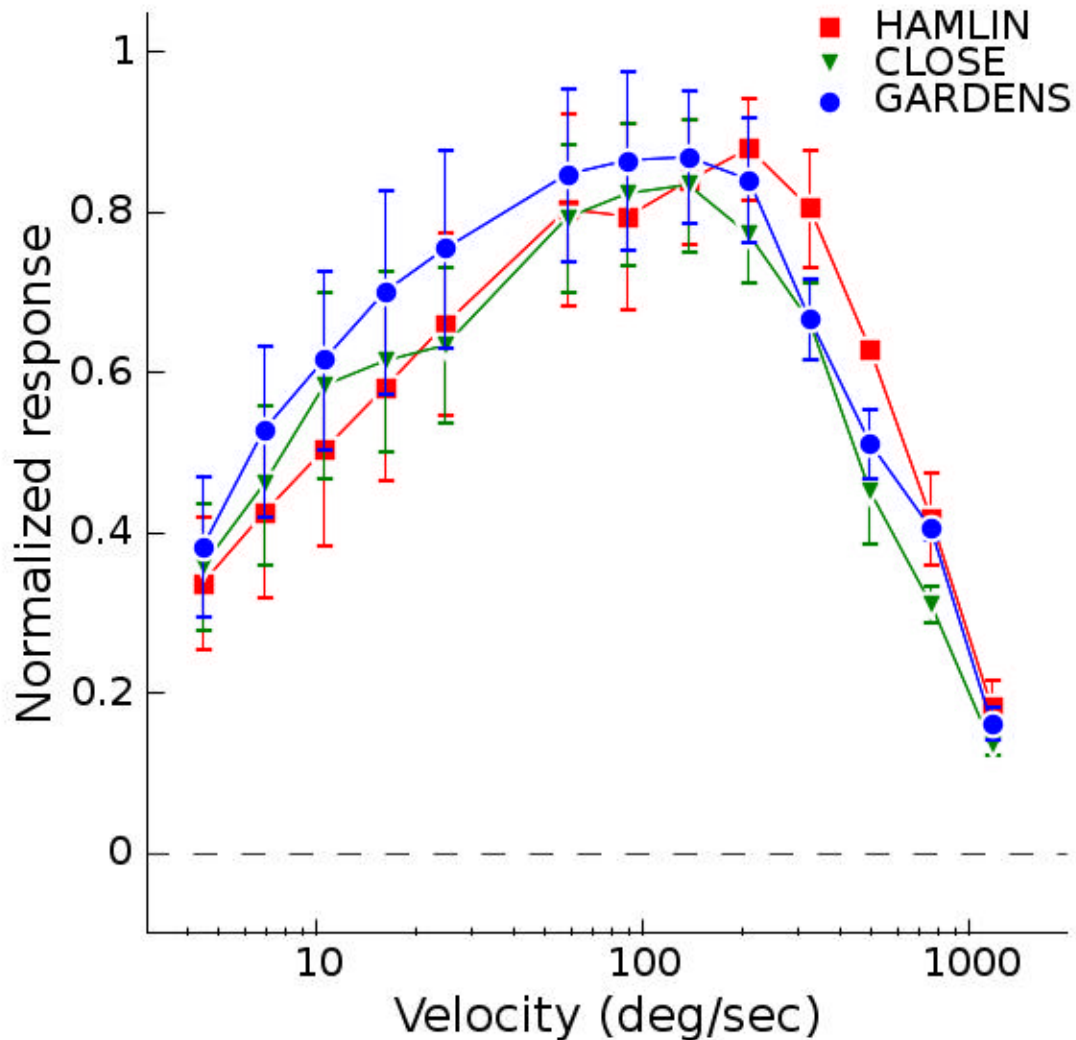


Fig. 1. Velocity tuning curves obtained from HSNE cells in the hoverfly *Eristalis tenax*. The HSNE is an identified HS neuron in *Eristalis* with a dorsal-equatorial receptive field. Normalized membrane potential (relative to rest) is shown for steady-state response to simulated yaw rotation with panoramic natural imagery. Experimental preparation and stimulus display are as described in Harris *et al.* (2000). The three images used as stimuli, identified by the labels at top, vary significantly in global contrast (these three images are depicted in Fig. 5 below, and simulated velocity tuning curves obtained with them from an array of correlational motion detectors is shown in Fig. 7). Each datum plotted represents an average of data from four male flies, for measurements made at four different phases of rotation 90° apart, averaged together over the period 0.2s – 1.0s following onset of motion, and normalized for each cell by the mean of the three maximum responses seen over all images and phases of motion for that animal. This normalization eliminates variability due to differences in scale of the curves / ‘quality of recording’ between animals. Error bars indicate ± 1 standard error.

With this study we examine the effects of a number of known and hypothesized features of visual motion processing that may contribute to velocity constancy. These are integrated into an overall model for wide-field motion sensing, from the compound eye to

the level of the tangential cell. Individual components of the model were chosen to capture the general characteristics of the biological elements in a computationally tractable form, rather than to account for their detailed neurophysiological properties. The model in a variety of configurations serves as the basis for a series of time-domain simulations, in which the input stimuli consist of the same animated natural imagery used in electrophysiological experiments.

2 The wide-field motion detection pathway

The compound eye in insects comprises a hexagonal array of *ommatidia*, which contain an individual lenslet, a lightguide structure, and a set of photoreceptor cells. After transduction in the retina, visual processing proceeds through three successive optical ganglia, in order, the lamina, medulla, and lobula. The wide-field tangential neurons in dipterans reside in a specialized portion of the lobula called the lobula plate. The lamina and medulla retain a retinotopic architecture, with one visual processing unit per ommatidium. Each unit operates primarily on a signal derived from a set of photoreceptors that view the same location in visual space. Lateral spatial interactions are also mediated by cells whose processes span multiple processing units. The local retinotopic architecture is for the most part lost in the lobula, where neurons receive inputs from fibers originating in broad regions of the medulla (or in some cases, the lamina) (Douglass and Strausfeld, 1995). The physiology of this pathway has been reviewed in detail elsewhere (Clifford and Ibbotson, 2003; Egelhaaf and Borst, 1993; Kirschfeld, 1972); we here review only salient aspects that relate to our modeling efforts.

2.1 Early vision

Visual processing begins with the compound eye optics, which are typically diffraction-limited and blur (i.e., spatially lowpass-filter) the image that appears on the insect retina to an extent that is well-matched to the spatial sampling period of the photoreceptors, preventing undersampling (Snyder, 1979; Snyder *et al.*, 1977). Phototransduction is nonlinear, with receptor membrane potential dependent on luminance in a roughly logarithmic manner. This dependence holds about an operating point that is adapted as a functional of the luminance history (van Hateren and Snippe, 2001; van Hateren, 1997). Fly photoreceptors are much faster in their temporal response properties than vertebrate receptors, but nonetheless have a lowpass characteristic with corner frequencies in the range 40Hz–70Hz in diurnal insects (Laughlin and Weckström, 1993). Insects generally possess color vision, but evidence suggests the pathways involved in motion detection are monochromatic (Srinivasan and Guy, 1990).

In the first optical ganglion, the lamina monopolar cells (LMCs) are thought to reside in the visual motion detection pathway, and they display a bandpass temporal characteristic relative to front-end luminance signals (James, 1992; Srinivasan *et al.*, 1982), with low-frequency rolloff below a few Hz.

2.2 Elementary motion detection

The operations of elementary motion detection are believed to take place primarily in the medulla. An early and still influential model for elementary motion detection, the correlational EMD, was formulated by Hassenstein and Reichardt (1956) based on

behavioral evidence, although its predictions are also consistent with many aspects of the physiology of motion sensitive neurons in the lobula and lobula plate. This model is based on a *correlation* of the signal associated with one visual processing unit with the delayed signal from a neighboring unit, as depicted schematically in Fig. 2.

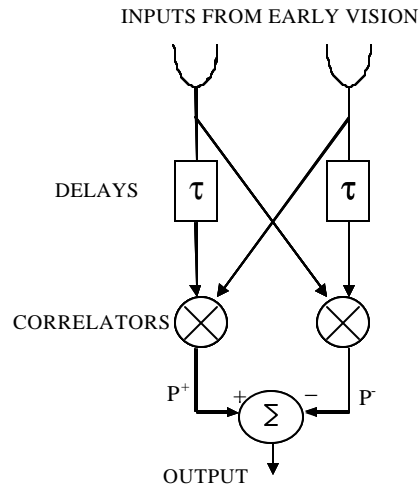


Fig. 2. The Hassenstein-Reichardt or correlational elementary motion detector. Inputs are from adjacent or nearby ommatidia in retinotopic space. This example is tuned to left-to-right motion: the output P^+ of the left correlator is in the mean greater than the output P^- of the right correlator for such motion.

The delay operator in the EMD is usually modeled with the phase delay of a lowpass filter, and the correlator with a multiplication. The final opponent stage in Fig. 2, which takes the difference of two mirror-image correlator outputs, enhances the directional properties and rejection of temporal contrast not due to motion. The correlational EMD produces a motion-related output without computing derivatives (a process that would amplify noise). Evidence suggests that EMDs are formed between at least nearest neighbors and next-nearest neighbors on the hexagonal lattice (Buchner, 1976), and are thus aligned with various directions in visual space.

The output of the correlational EMD in response to a moving visual scene is typically unsteady, with transients generated in response to the passage of edges or contrast gradients. In the *mean*, it is a function of velocity of a moving stimulus, although this dependence is not monotonic, and there is also strong dependence on spatial structure and contrast of the stimulus. In spite of its ambiguities as a motion sensor, theory suggests that this type of detector is inherently well suited to tasks that might be limited by noise (Potters & Bialek, 1994).

2.3 Spatial integration by tangential neurons

Tangential cells are large neurons with extensive dendritic arborizations, and they are believed to integrate signals from elementary motion detectors over wide areas of the visual field (Krapp *et al.*, 1998; Egelhaaf *et al.*, 1989; Franceschini *et al.*, 1989). This naturally provides a signal averaging effect with respect to the unsteady outputs of individual EMDs. Variable weighting of these inputs, by means of differing synaptic efficacies, is believed to be responsible for the variations in absolute sensitivity and

direction preference of the tangential cell to visual motion across its receptive field. Van Hateren (1990) has shown how weighting and summing the outputs of local EMDs aligned with different interommatidial axes can yield a motion detector with an arbitrary preferred-direction response, and that three nearest-neighbor EMDs in a hexagonal system can give a detector with near-ideal, cosine-like directional sensitivity.

Evidence suggests, however, that the integration of inputs by tangential cells cannot be represented as a simple weighted sum. In particular, when the visual system is subject to moving patterns with very low contrast or subtending only a small part of the visual field of a tangential cell, its response varies as the contrast or stimulus area is increased, as would be expected if the processing were linear. However, as stimulus conditions approach typical contrasts and full-field motion, this dependence vanishes and the response saturates, although at a value that still depends on the velocity and spatial structure of the stimulus (Haag *et al.*, 1992; Hausen, 1982). This effect has been modeled as a form of gain control induced by shunting inhibition at the inputs to a tangential neuron, mediated by a second motion-sensitive ‘pool cell’ (Poggio *et al.*, 1981). However, Borst *et al.* (1995) have shown how it could arise from synaptically-mediated ion conductances in the cell membrane of the tangential neuron itself. This model lacks any voltage-gated membrane conductances, and in its simplest form is a single electrical compartment. Its operation may be analogized as a voltage division in which the membrane conductance to an ion species with a depolarizing reversal potential is mediated by one class of inputs, conductance to an ion species with a hyperpolarizing reversal potential is mediated by a second class, and a ‘fixed’ conductance to one or more ion species is also present and determines the cell’s resting membrane potential.

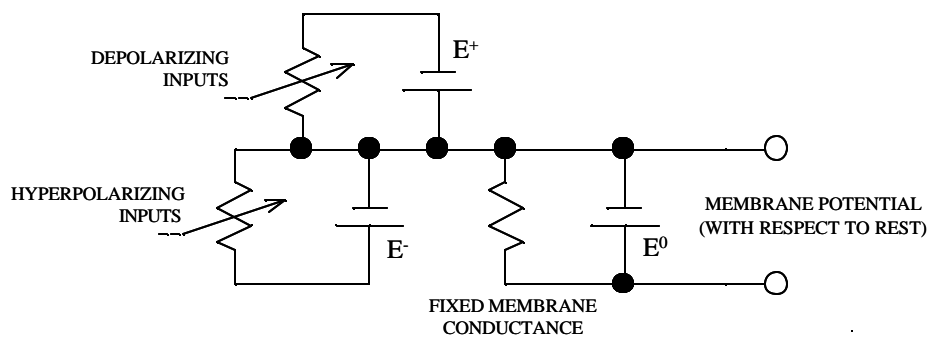


Fig. 3. Schematic diagram illustrating the electrical principle of operation of the single-compartment version of the ‘gain control’ model of Borst *et al.* (1995). Membrane potential in Eqn. (1) is measured with respect to the resting state, which is determined by a fixed membrane conductance to one or more ion species with an effective net reversal potential of E^0 . Depolarizing inputs mediate membrane conductance to a ion species with reversal potential E^+ that is positive with respect to rest, and hyperpolarizing inputs mediate conductance to a ion species with reversal potential E^- that is negative with respect to rest. When these synaptically-mediated conductances dominate the fixed conductance, the membrane potential is determined only by the relative strengths of the depolarizing and hyperpolarizing inputs.

It is supposed that depolarizing and hyperpolarizing classes of inputs correspond respectively to the outputs P^+ and P^- of complementary pairs of correlators as represented in the EMD schema in Fig. 2. Let the index i ($i=1\dots n$) designate the principal directions or axes with which EMDs are aligned, and index j ($j=1\dots m$) indicate position of the

EMDs within the receptive field area of a tangential cell. Then the wide-field neuron computation with the Borst ‘gain control’ model can be expressed in the form:

$$E = \sum_{i,j} (G_{ij}^+ E^+ - G_{ij}^- E^-) / (\sum_{i,j} (G_{ij}^+ + G_{ij}^-) + G_0), \quad (1)$$

where the membrane potential E and the depolarizing and hyperpolarizing reversal potentials E^+ and E^- , respectively, are measured with respect to rest, where G_{ij}^+ and G_{ij}^- are the synaptically-mediated conductances, and G_0 is the ‘fixed’ membrane conductance. Saturating behavior occurs when the variable conductances are large enough to dominate the fixed conductance: the response becomes dependent only on relative and not absolute magnitudes of the G_{ij}^+ and G_{ij}^- . The G_{ij}^+ and G_{ij}^- reflect both synaptic efficacy or weighting, and the strength of output from individual correlators. In this way, a saturating and strongly directional response is obtained from the weakly directional outputs of individual correlators. Borst and colleagues refer to this mechanism as ‘gain control’ (although in this case it is a static nonlinearity and not gain control in the engineering sense). The single-compartment version of the model has been used in subsequent work (Kern *et al.*, 2001; Harrison and Koch, 2000); a more realistic multicompartment version was shown to display qualitatively similar behavior (Borst *et al.*, 1995).

2.4 Motion adaptation

When the visual system of an intact animal is exposed to strong motion stimuli and subsequently probed with small impulses or steps in velocity, the response of a tangential cell to these test stimuli is reduced relative to that of an initially unstimulated cell (Harris *et al.*, 2000; Clifford and Langley, 1996; de Ruyter van Steveninck *et al.*, 1986; Maddess and Laughlin, 1985). This phenomenon has been termed motion adaptation. While in the past it has been suggested that its mechanism is a reduction in the time constant of the delay operator in a correlational EMD (Clifford *et al.*, 1997; Borst and Egelhaaf, 1987; de Ruyter van Steveninck *et al.*, 1986), recent work has produced strong evidence that this is not the case, and that it is actually due to a reduction in gain somewhere in the prior signal processing pathway, as well as a shift in the resting membrane potential of the cell when the adapting motion is in a direction that causes net excitation of the cell (Harris *et al.*, 2000; Harris *et al.*, 1999). The gain reduction is elicited by motion in any direction, even when it causes little response or net inhibition in the cell. It has also been shown to be spatially local (i.e., it occurs prior to integration of motion detector outputs) (Harris *et al.*, 2000) and to vary directly with contrast of the adapting stimulus when that contrast is above some minimal threshold (Straw and O’Carroll, unpublished observations). Dependence on contrast, i.e., relative rather than absolute variations in luminance in the moving scene, is presumably due to the properties of early vision.

The details of what drives the process of motion adaptation remain a subject of research. We herein model the gain reduction phenomenon as a form of local gain control at the front end of the EMD operation. Our simple approach is consistent with the characteristics mentioned in the previous paragraph, but it is by no means complete: other features of motion adaptation have come to light that it cannot explain. In particular, motion adaptation seems to occur on multiple and largely different time scales (Fairhall *et al.*, 2001), and appears to be more strongly recruited by motion signals than by purely temporal contrast such as flicker (Harris *et al.*, 2000). This suggests that it may be a

complex process with feedback of motion signals, or possibly adaptation at multiple stages in the motion processing pathway.

2.5 Velocity Constancy

Given a general picture of the wide-field motion detection pathway, we ask, what features may contribute to the experimentally-observed property of velocity constancy in the tangential cells? Part of the answer undoubtedly lies with the relative consistency of the spatial statistics of natural scenes themselves (Ruderman, 1994; Tolhurst *et al.*, 1992). The spatial power spectra of luminance along one-dimensional paths in natural images follow an approximate $1/f$ characteristic, where f is spatial frequency (Dror *et al.*, 2001; van Hateren, 1997). In addition to similarity between different scenes, this characteristic implies a certain self-similarity of natural imagery at different spatial scales (i.e., at different ranges from the eye), and it leads to a corresponding consistency in the response of the correlational EMD model: velocity tuning curves (mean output versus velocity of optic flow) obtained with different moving natural images tend to be very similar in shape (Dror *et al.*, 2001; see also Fig. 7 below). However, natural scenery can vary considerably in contrast, and due to the quadratic nature of the EMD correlator function, this results in large differences in the amplitude of velocity tuning curves obtained with different scenes. Kirschfeld (1991) has noted that this contrast dependence has implications for optomotor control as well as motion sensing *per se*. We expect that operations which in some way reduce the contrast dependence of signals in the processing path should reduce variations in response to differing natural scenery, and contribute toward velocity constancy. Two such operations are motion adaptation and nonlinear spatial integration by the tangential cells.

An additional such feature that is biologically plausible is the presence of saturating nonlinearities in the processing chain, which might arise simply due to the limited signaling ranges of the neural components in the pathway. Saturating nonlinearities have been included as elaborations in the correlational model for motion detection, and have been shown to improve reliability of velocity estimation (Dror *et al.*, 2001).

Finally, we consider the possibility that spatial highpass filtering may take place in the pathway somewhere prior to motion detection. The on-center, off-surround spatial opponency that implements highpass filtering is ubiquitous in neurobiology, and prior studies of predictive coding in the visual system (Srinivasan *et al.*, 1982) and on LMCs (van Hateren, 1992) find evidence of suppression of low spatial frequencies in LMC responses. We have also found some evidence for such filtering in the response characteristics of wide-field motion-sensitive neurons in diurnal sphingid moths (O'Carroll, unpublished observations), although as yet not in tangential neurons in flies. Simple correlational EMDs can give excellent fits to spatial frequency tuning data (obtained with sinusoidal images) from fly tangential cells without the inclusion of any spatial high-pass filtering on the inputs (O'Carroll, unpublished observations), but the influence of unknown levels of compressive non-linearity and other nonlinear characteristics of the real system may also affect the sinusoidal tuning data. The inclusion of spatial highpass filtering is further motivated by the observation that the natural scenes we have used in our studies, in spite of overall similarity of their spatial power spectra, still show individual variations in spatial structure and these are more prevalent in the low (and higher-power) frequency components. Reducing the effect of these variations by

spatial highpass filtering might contribute toward velocity constancy. We discuss this in more detail in the sections that follow.

3 A system-level computational model and simulations

We assembled a computational model for wide-field visual motion processing from the compound eye to the level of the tangential cell, with the aim of providing a high-level framework for simulation of this pathway. The basic version of this model includes a model for photoreceptor nonlinearity, temporal filters to mimic the dynamic properties of cells in the retina and lamina, correlational EMDs, and summation of EMD outputs by a model tangential cell. To this basic model we added features that we expected to contribute to velocity constancy in the tangential neuron. These include: 1) spatial highpass filtering in early vision; 2) motion adaptation modeled as contrast gain control prior to the EMD; 3) and 4), saturating nonlinearities applied to the outputs of early vision or the outputs of correlators in the EMD model, respectively; and 5) nonlinear integration by the model tangential cell, in the form of the ‘gain control’ model of Borst *et al.* (1995).

We performed time-domain simulations to investigate the contribution of these features toward the ideal of velocity constancy. In these simulations, data representing animated natural imagery were processed according to the model in its various configurations. Five different natural images, varying substantially in contrast and spatial structure, formed the data set. Motion in all cases consisted of uniform horizontal translation, to which the model tangential cell was tuned. The model was implemented so that each of the test features detailed in the previous paragraph could be optionally included in the simulations in order to study its effect. Velocity constancy was evaluated by examining the scatter in the steady-state ‘tangential cell’ responses over the set of five images, for a particular model configuration and speed of image motion.

3.1 The model

Visual transduction, early visual processing, and elementary motion detection were specified as taking place on a hexagonal lattice with one of the interommatidial axes oriented in the vertical (latitudinal) direction. This arrangement mimics the geometry of the compound eye and the retinotopic distribution of visual processing units in the lamina and medulla. Following these stages of processing, a model tangential cell integrated the EMD outputs. This processing is depicted graphically in the flow chart in Fig. 4.

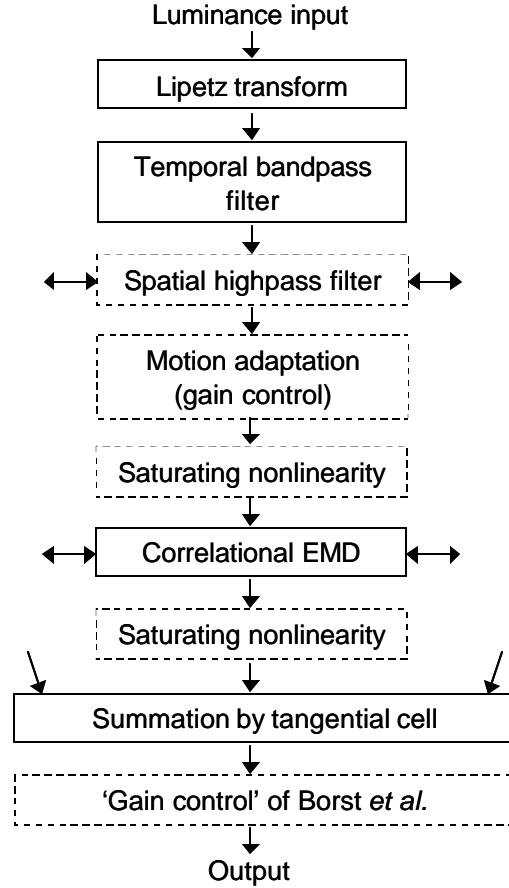


Fig. 4. Summary of the model used for wide-field motion processing. Up to the level of the tangential cell, the flowchart depicts a single visual processing unit. Solid text boxes indicate elements of the basic model; dashed boxes represent additional test features evaluated in individual simulations. Horizontal arrows indicate lateral spatial interactions.

The model elements in order are as follows. The geometrical optics of the model were scaled to reflect the approximate characteristics of the eyes of large, visually acute dipterans. The equivalent interommatidial angle was in the range $1.20^\circ - 1.25^\circ$, and ommatidial geometry was fixed across the model receptive field. Compound eye optics were modeled by blurring of relatively high-resolution input imagery by a Gaussian modulation transfer function (Götz, 1964) F of the form

$$F(\mathbf{f}) = \exp[-2.77\mathbf{f}^2 / (\Delta\mathbf{r})^2], \quad (2)$$

where \mathbf{f} is angular deviation from the optical centerline of an individual ommatidium, and the parameter $\Delta\mathbf{r}$ was set to 1.64° .

The early vision model includes a Lipetz (1971), or Naka-Rushton (Naka and Rushton, 1966) transformation followed by a temporal bandpass filter. The Lipetz transformation is specified by the equation

$$U = I^a / (I^a + I_0^a), \quad (3)$$

where I is the input intensity, I_0 a parameter defining mid-response level, a is an exponent between 0.5 and 1, and U is the output. This equation gives log-linearity of response over roughly a decade of intensities about the mid-response level. The exponent a was set to the value 0.7.

Although a role for photoreceptor adaptation in velocity constancy cannot be ruled out *a priori*, this phenomenon was not modeled in our simulations for several practical reasons. The luminance data used in the simulations are derived from photographic images, and setting the exposure on a camera performs much the same function as photoreceptor adaptation in the biological retina (although in a global rather than local sense). The data are also relatively limited in dynamic range. Due to these considerations, we rather set the mid-response level I_0 in (3) for each individual image to a fixed value corresponding to an estimate of the geometric mean of the luminance over that image.

The bandpass filter in early vision was a linear operator with transfer function (Laplace transform) $\tau_{HS}/[(\tau_{HS}+1)(\tau_{LS}+1)]$, where s is the Laplace variable, and where τ_H was set to 400ms and τ_L to 8ms, corresponding to corner frequencies of about 0.4Hz and 20Hz, respectively.

Strong spatial highpass filtering, with a space constant corresponding to one interommatidial separation, was used when this feature was included. The filter was implemented by applying a spatial lowpass stage and subtracting its output from the local signal. The lowpass filter was based on a linear, continuous diffusive model (approximating electrotonic spread through a neural network with dense gap junctions) with an exponentially decaying spatial impulse function

$$V(r) = k \cdot \exp[-r / I], \quad (4)$$

where k is a normalization constant, r is the Euclidean distance from the application of the impulse measured in units corresponding to the interommatidial separation, and I is the space constant. The scaling of the local signal and the constant k were chosen so that application of a unit impulse resulted in a unit-magnitude output from the highpass filter.

Motion adaptation (when included) was modeled as local gain control applied to early vision outputs. The mean absolute deviation (MAD) of each individual early vision output signal was estimated by full-wave rectification followed by a linear, first-order temporal lowpass filter (transfer function $1/(\tau_{AS}+1)$), with fixed but user-selectable time constant τ_A . The early vision signal was then divided by the MAD estimate. The time constant τ_A is a measure of the characteristic time scale of adaptation. This gain control model is similar to the approach of Kirschfeld (1991) except for the point of application of the variable gain.

Saturation of early vision signals (when included) was modeled with multiplication by a scaling factor followed by application of a hyperbolic tangent function. This saturating nonlinearity was applied following gain control in the cases when motion adaptation was present. (It would make little sense from a signal processing standpoint to apply gain control to a signal that is already significantly clipped, whereas it *is* sensible to regulate the amplitude of a signal in order to make full use of the dynamic range of a following stage which limits the amplitude of signals that it passes.) The scaling factor was chosen based on the degree of signal limiting observed with the input data used in simulations. A 'moderate' level of saturation, i.e., a characteristic between linear transformation and

binary thresholding, was the objective of this choice. A quantitative characterization of the criterion is given in the Appendix.

Elementary motion detection consisted of the correlational EMD depicted in Fig. 2. The delay operator was modeled as a linear first-order lowpass temporal filter (transfer function $1/(\tau_D s + 1)$), and the correlation operation as a multiplication. The time constant τ_D of the lowpass filter was set to 40ms. Only EMDs formed between nearest neighbors on the hexagonal lattice were included.

Saturation of the EMD correlator outputs (when included) was implemented in the same way as saturation of early vision signals.

All elements described to this point were assumed to be distributed one per visual processing unit, or in the case of the EMD, one per each neighboring pair of units. The next step in the processing chain is the integration of EMD output signals by a tangential cell. We implemented a single model tangential cell, sensitive to longitudinal motion across a receptive field of about 40° height (latitude) and 160° width (longitude). With the input imagery used, this element may be regarded as roughly equivalent to an equatorial HS cell. There was, however, no variation in absolute or directional sensitivity to motion across the receptive field, as there is in the biological HS cell: all EMDs with an orientation of 60° with respect to vertical, and 120° with respect to vertical, were given equal weight as inputs to the ‘HS cell’, in order to form a motion detector with maximum and equal sensitivity to longitudinal motion (van Hateren, 1990), everywhere in its receptive field.

We specify HS cell response in terms of the individual correlator outputs from the EMDs. As in (1), let the index i ($i=1,2$) designate the directions or axes with which input EMDs are aligned, and index j indicate position of an EMD within the receptive field area. The basic HS cell simply sums EMD outputs over its receptive field:

$$Y = \sum_{i,j} (P_{ij}^+ - P_{ij}^-), \quad (5)$$

where Y is the cell response, and P_{ij}^+ and P_{ij}^- represent the outputs of complementary correlators in the EMD, as in Fig. 2. The response of the HS cell with the ‘gain control’ model of Borst *et al.* (when this feature is included) is given by

$$Z = \sum_{i,j} (P_{ij}^+ - P_{ij}^-) / \sum_{i,j} (P_{ij}^+ + P_{ij}^-), \quad (6)$$

where Z is the cell response. This formulation tacitly assumes that the products of synaptic efficacies and reversal potential magnitudes are equal for excitatory and inhibitory synapses, and that synaptically modulated membrane conductances dominate ‘fixed’ conductances. We make the former assumption because we are interested not in predicting the numerical values of membrane potential in a real HS cell, but only in evaluating the contribution of the ‘gain control’ model in reducing the variance in responses to different scenes. The second assumption is justified because we are considering full-field motion of imagery at natural contrasts, under which conditions a tangential cell would be expected to be in the saturated or ‘gain control’ regime.

3.2 Simulation environment

We performed simulations with SPICE (Simulation Program with Integrated Circuit Emphasis), a tool for circuit simulations that is optimized for integration of stiff nonlinear differential equations. Discrete elements in SPICE are described by constitutive relations

between electrical quantities at interconnection points or terminals, and governing equations for interconnected systems of elements are derived by application of Kirchhoff's current law (conservation of charge). For time domain simulations, SPICE integrates the governing equations numerically to obtain the evolution of the state of dynamical systems.

SPICE includes model elements for various electronic devices, but also permits the implementation of abstract elements in the form of voltage and current sources whose states may be specified by equations written in terms of terminal voltages throughout the system. Such elements were used for the implementation of nonlinear processing stages in the model. Linear stages were implemented with resistors, capacitors, and linear sources.

External inputs to a system in time-domain simulations may be made by means of piecewise-linear sources. Sampled time-series of input data are supplied by the user, and SPICE interpolates the data between sample points as required during integration. This interpolation is nominally linear, but a rounding feature may be invoked to obtain cubic spline fits between data points. We used piecewise-linear sources with the rounding feature invoked, to supply the visual input in all simulations performed.

3.3 *Input data*

Five different high-resolution panoramic images of predominantly natural scenes were used to derive input data. These scenes all comprise known habitat for several species of dipterans. A series of 12 color photographs of each scene was taken at 30° intervals with a digital camera rotated longitudinally about the nodal point of the lens, and the corresponding panorama was stitched together using Apple Quicktime VR Authoring Studio software. The vertical extent of each is about 54°, with the horizon at center. Final sizes are 2048 pixels by 308 to 320 pixels. The eight-bit green values were extracted from the images as most closely matching the spectral range passed by the monochromatic receptors believed to be involved in motion vision in the flies (Srinivasan and Guy, 1990). Three of the panoramic images are shown below in Fig. 5.



Fig. 5. Three of the five panoramic images used as sources for input data in simulations. Eight-bit green values from original color photographs are depicted. The top image, assigned the identifier '*hamlin*', has the highest global contrast; the center image, '*close*', has a mid-range global contrast; and the bottom image, '*gardens*', has the lowest contrast in the set.

Data from these images were resampled onto a hexagonal grid corresponding to the simulated ommatidial array described below in Section 3.4, and at the same time blurred by discrete spatial convolution on the original rectangular pixel lattice with the modulation transfer function in (2). Interommatidial angle was set to seven pixels in the vertical direction, corresponding to 1.23° , and the horizontal pitch of columns of ommatidia was set to six pixels (leading to a horizontal/vertical aspect ratio of about 0.99). The convolution was carried out over a finite square region of support of extent $3 * \mathbf{Dr} / \sqrt{2.77}$ in the vertical and horizontal dimensions.

Animation was achieved by shifting the resampling centers rightward by a fixed amount for each simulated time step, and repeating the resampling/blurring process. The horizontal step size was fixed at two pixels, insuring that the processed data were oversampled. Two full rotations through each panoramic image were performed. A time series of processed data for each simulated ommatidium in the array was written out in the format of a SPICE statement defining a piecewise-linear source. The sample period was specified as a parameter, rather than written numerically. In this way, the speed of motion in any particular simulation could be set as desired by assigning a value to this parameter at execution time.

Although in the biological system any strong spatial highpass filtering is presumed to occur after the bandpass temporal filtering of early vision, in our model both of these operators are linear, and thus the order in which they are applied does not affect the computation. With this in mind, we prepared a second set of input data in which the initial static transforms – blurring, the Lipetz transformation, and highpass filtering – were performed during the processing and animation of the images. These data were intended for use in simulations that included spatial highpass filtering, with the Lipetz transform removed from the processing chain in the simulation itself. This allowed us to maintain the computationally efficient simulation approach described in Section 3.4 below. The highpass filter was implemented by discrete spatial convolution with the lowpass kernel in (4) on the hexagonal lattice with support over nearest through third-nearest neighbors, followed by subtraction from the (scaled) local signal.

During processing of each image, we also computed an estimate of the geometric mean of the luminance values (range [0,255]) over the entire panorama for purposes of setting the mid-level response parameter in (3). Zero-value pixels (which comprised about 2% of the pixels in the darkest image) were replaced with the value 0.25 for this computation.

We also computed spatial power spectra for luminance in the horizontal direction in each image. An FFT was performed on each row, and the resultant power spectra averaged over all rows. Log slopes obtained by least-squares fits were in the range -0.99 to -1.30 . In spite of overall similarity of the spatial spectra, differences in the structure of the scenes were evident and reflected in the relative distribution of power among the bands. We computed the relative proportion of ac power in each band for each image, and then the mean and standard deviation of these quantities over all five images. The standard deviation relative to the mean, shown below in Fig. 6, gives a measure of how

structural dissimilarity between the scenes is distributed according to spatial frequency. The largest variations are seen at lower frequencies. The very lowest frequency components make little contribution to the output of the correlational EMD because there is negligible phase difference between two ommatidia for such frequencies, but frequencies between 0.05 cycles/degree and 0.15 cycles/degree would certainly have an influence, given the interommatidial spacing we use in our model. These data support the notion that spatial highpass filtering might contribute to velocity constancy by reducing the amplitude of low-frequency components that display the greatest interscene variance.

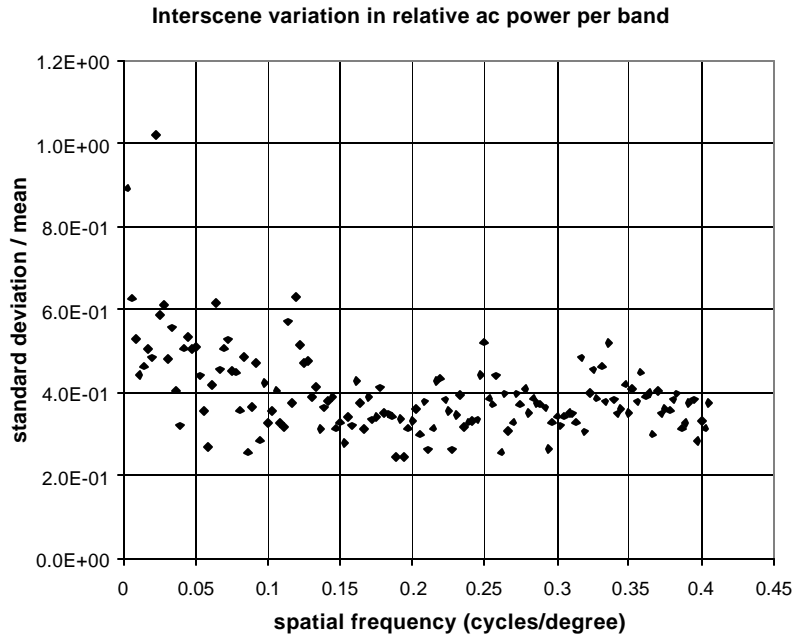


Fig. 6. Interscene variation in the distribution of ac power per band, from power spectra obtained by FFTs performed on rows of individual images. The spatial frequencies shown range up to the Nyquist frequency for spatial sampling in the ommatidial array. Standard deviation and mean are computed over the sample of five images used in the study.

3.4 Simulation approach

Our simulation approach involves the characterization of the steady-state response of a model HS cell to uniform longitudinal translation of imagery across its receptive field. Due to the nature of this stimulus, horizontal displacement of the image and time are equivalent as independent variables. We were able to compute the model HS cell responses based on simulation of an array of visual processing units that is much smaller than the full HS cell receptive field, by taking advantage of the fact that the time series of luminance for any two ommatidia at the same latitude are equivalent except for a time shift.

The simulated array consists of three retinotopically-adjacent columns of 33 visual processing units each. The center column in this hexagonal array is offset downward by half of the interommatidial separation, relative to the left and right columns. EMDs oriented with the 60° and 120° axes are formed between adjacent units in the left and

center columns, and adjacent units in the center and right columns. A total of 65 EMDs of each orientation are present in this arrangement. Three columns of ommatidia and two columns of EMDs are necessary because adjacent ommatidia do not see equivalent time series of inputs due to the vertical offset between alternate columns in the hexagonal architecture.

During simulations, the difference between complementary EMD correlator outputs (the right-hand side of (5), or the numerator in (6)) is summed over the array, as is the sum of the correlator outputs (the denominator in (6)). These two quantities are written out by the simulation program at time intervals corresponding to exactly two columns' worth of horizontal displacement. Any set of consecutive values in this series may be regarded not just as a time series generated by a single array, but also as *instantaneous* values obtained from a contiguous set of such arrays covering a broader receptive field, and viewing the same image moving at the same speed. (This equivalence applies to steady-state responses, in which transients due to the onset of motion have died out.)

Two full rotations of each image were performed during each simulation. The difference and sum outputs were transferred to a spreadsheet, and at each time step during the final rotation, the current and 74 prior values of each were summed. In this way, estimates were formed for the corresponding quantities for an HS cell with a receptive field 150 columns wide, over one full rotation of the image. Based on each pair of values, the quotient Z in (6) was also computed. The difference term alone represents the output Y of the basic HS cell model, whereas Z is the output of the HS cell including the Borst 'gain control' model. To characterize the steady-state response, the average and standard deviation of each were computed over the entire rotation, and are reported as results in the following section.

To exclude transient effects due to the onset of motion, we discarded the initial simulation output data, corresponding to the first rotation of the image less the HS cell receptive field width. In all simulations we limited the speed of motion such that the time covered by the discarded data was more than twice the longest time constant in the system (that is, the 400ms time constant associated with the temporal highpass filter in early vision). The maximum speed that could be simulated under this restriction was about 252°/s.

For model configurations that did not include spatial highpass filtering, all operations associated with the neural pathway were performed in individual simulations. When the highpass filtering was included, a modified simulation program was used in which the Lipetz transform was omitted and processing began with bandpass temporal filtering.

For each configuration of the model, simulations were performed with several images. Velocity tuning curves for the basic model were obtained for the three images depicted in Fig. 5, which are those with the highest, lowest, and a mid-level global contrast. However, the bulk of the simulations were performed at a single speed (50°/s) near the velocity optimum for the basic model, and for all five images, with the aforementioned aim of evaluating scatter in the response of particular model configurations over the range of natural imagery in the dataset. Scatter was quantified as standard deviation divided by mean over the five images. A series of simulations was run in which motion adaptation was included, and in which the time constant τ_A was set to four different values (32ms, 80ms, 200ms, and 500ms); the value which minimized scatter (200ms) was used in subsequent simulations in which adaptation was included. All possible

combinations of the test features were evaluated, except that only one of the two saturating nonlinearities was included in any given run.

4 Results

Following in Fig. 7 are velocity tuning curves for the basic model, for the images *hamlin*, *close*, and *gardens*. The strong dependence of EMD output on image contrast is evident in the large differences in the scale of the individual curves. However, the similarity in their shape and in the velocity optima (which is clear if each is normalized by its maximum datum) reflects the general similarity in spatial structure of the natural scenes that form the input stimuli.

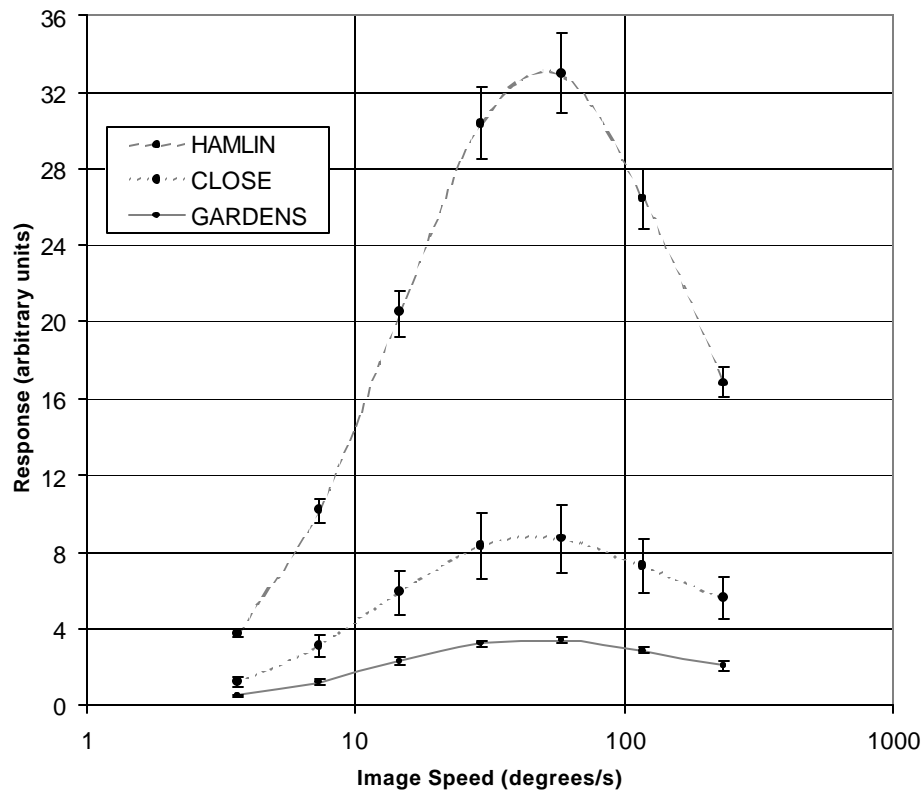


Fig. 7. Velocity tuning curves for the basic model, for the animated images *hamlin*, *close*, and *gardens* depicted in Fig. 5. Data points represent mean values of HS cell output over one image rotation, and error bars indicate ± 1 standard deviation of the instantaneous output value over the course of the simulation, and give an indication of output variation due to local variations in scene structure.

Each of the test features was found to result in reduction in inter-scene scatter at the test speed of 50°/s, relative to the basic model. Depicted below in Fig. 8 are results from simulations including each feature individually, as well as from the basic model:

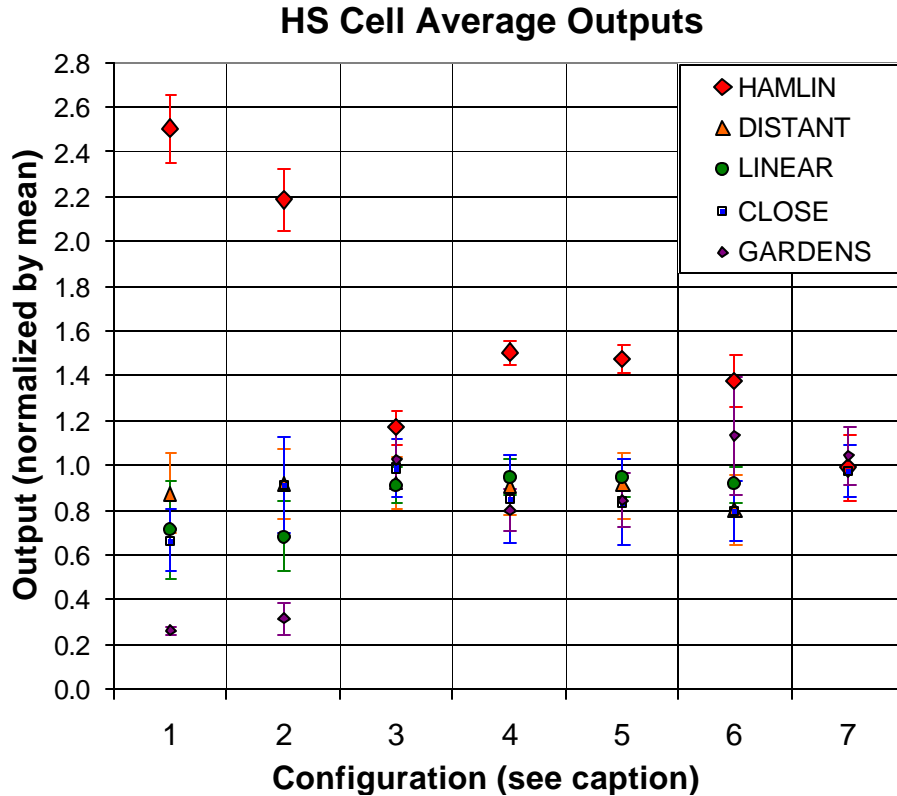


Fig. 8. Variability of responses between different animated scenes of the model in various configurations, at the near-optimal speed of 50°/s. In each column of the graph are displayed responses (average over one image rotation) to all five images for a particular model configuration, normalized by the mean over all five. Error bars for simulation results indicate ± 1 standard deviation of the instantaneous output value over the course of the simulation, also normalized by the mean over all five images. Configurations include each of the test features individually: 1. basic model; 2. with spatial highpass filtering; 3. with motion adaptation (gain control) with time constant $\tau_A = 200\text{ms}$; 4. with saturation of early vision signals; 5. with saturation of EMD correlator outputs; 6. with the Borst ‘gain control’ model in the HS cell. Column 7 shows mean responses of biological HS cells in the hoverfly *Eristalis tenax* to yaw stimuli with the three images depicted in Fig. 5, at a speed of 58°/s.

Also shown in Fig. 8 are three data obtained from biological HSNE cells in the hoverfly *Eristalis tenax*, in response to the images *hamlin*, *close*, and *gardens* in simulated yaw rotation at 58°/s. These are the same data that appear in Fig. 1, except renormalized by their mean. The error bars in Fig. 8 give an indication of output variation due to local variations in scene structure (as well as random noise in the case of the biological data). The inter-scene scatter for the three scenes is smaller in the biological data than for any of the model configurations included in the figure.

Following in Fig. 9 are shown the results for the complete set of simulations performed to evaluate contribution of the test features toward velocity constancy. The performance of each model configuration is characterized by the single interscene scatter value.

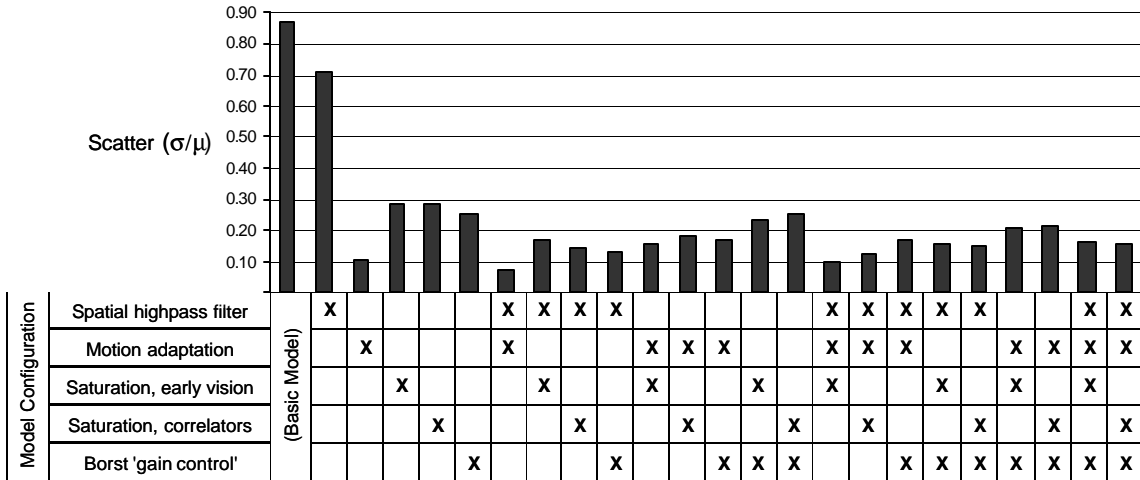


Fig. 9. Summary of results from all simulations showing inter-scene scatter in model response to the five natural test images used in simulations. Headings at left indicate the test features, and an ‘X’ indicates the inclusion of a feature in the model configuration represented by each column. For runs including motion adaptation, the time constant $\tau_A = 200\text{ms}$. Scatter is measured as standard deviation divided by mean over the responses to the five-image set, where each response is the average output over one image rotation.

Several conclusions can be drawn from these data. The simple gain-control-based motion adaptation model provides the greatest reduction in scatter; the four lowest values are achieved by configurations that include it. Spatial highpass filtering by itself does not greatly improve scatter when compared to the basic model, but in combination with any one of the nonlinear features, it significantly improves performance relative to the feature by itself. The same is true when multiple nonlinear features are present, in all cases but one. Interestingly, combining two or more nonlinear features does not generally result in reduced scatter. In all such cases but one, better performance can be obtained with some combination of fewer of the features.

5 Discussion and conclusions

Velocity constancy, which in a wide-field motion-detecting neuron we define as dependence of mean response on velocity of optic flow, in combination with invariance with respect to differing natural scenes, is closely approximated by the tangential neurons in the lobula plate of various fly species. The responses of these cells depend upon the pattern and overall scaling, or speed, of optic flow that is present on the retina, but they appear to be capable of rejecting the influence of variations in contrast and spatial structure that occur in natural imagery. The relative consistency of the spatial statistics of natural scenes may contribute to this capability, but because variations in contrast and differences in structure are still present, the processing that takes place in the visual pathway must play a major role. We have tested a number of plausible or established features of this processing, which we expect to reduce inter-scene scatter in wide-field motion detection. These include spatial highpass filtering, motion adaptation (modeled as gain control), saturating nonlinearities in the signal path, and nonlinear integration by a tangential cell analog, as modeled by Borst *et al.* (1995). Simulations of the visual

pathway from compound eye to the tangential cell were run, in which these features were included singly and in most possible combinations, and in which the input data were derived from a set of animated natural images. The primary conclusion that we draw from the results is that, while all of the test features reduce inter-scene scatter relative to the most basic version of the model, in no case does the model approach velocity constancy as closely as does the biological system. We also find that the only linear feature (spatial highpass filtering) significantly improves the performance of the model in combination with one or more of the nonlinear features, although it results in little reduction in interscene scatter by itself. This we suggest is due to reduction of the amplitude of low spatial frequency components of natural imagery, which tend to have large interscene variance. Finally, we note that combining multiple nonlinear features does not generally improve performance.

We speculate that motion adaptation may play an important role in the velocity constancy observed in biological tangential cells, and that our simple model for gain modulation is not adequate to explain its full effects. From a signal processing standpoint, it seems clear that velocity constancy is a desirable property for a visual motion detection system, as its primary effect is elimination of dependence on parameters of visual scenes that are irrelevant with respect to motion. If this confers clearer and less ambiguous information about self-motion to an organism, it is conceivable that evolutionary pressures might have driven the development of motion adaptation (as well as other features of wide field motion detection) in a manner that subserves velocity constancy.

Acknowledgements: This work was supported by US Air Force SBIR contract F08630-02-C-0013 and by US Air Force IRI grant F62562-01-P-0158. A. Straw was supported by a fellowship from the Howard Hughes Medical Institute. Data on velocity constancy were contributed in part by T. Rainsford. The authors thank T. Bartolac for data processing and for comments on the manuscript.

References

- Borst A, Egelhaaf M, Haag J (1995) Mechanisms of dendritic integration underlying gain control in fly motion-sensitive neurons. *Journal of Computational Neuroscience* 2:5-18.
- Buchner E (1976) Elementary movement detectors in an insect visual system. *Biological Cybernetics* 24:85-101.
- Clifford CWG, Ibbotson MR (2003) Fundamental mechanisms of visual motion detection: models, cells and functions. *Progress in Neurobiology* 68:409-437.
- Clifford CWG, Langley K (1996) Psychophysics of motion adaptation parallels insect electrophysiology. *Current Biology* 6:1340-1342.
- de Ruyter van Steveninck R, Zaagman WH, Mastebroek HAK (1986) Adaptation of transient responses of a movement-sensitive neuron in the visual system of the blowfly *Calliphora erythrocephala*. *Biological Cybernetics* 54:223-236.
- Douglass JK, Strausfeld N (1995) Visual motion detection circuits in flies: Peripheral motion computation by identified small-field retinotopic neurons. *Journal of Neuroscience* 15:5596-5611.
- Dror RO, O'Carroll DC, Laughlin SB (2001) Accuracy of velocity estimation by Reichardt correlators. *Journal of the Optical Society of America A* 18:241-252.
- Egelhaaf M, Borst A (1993) A look into the cockpit of the fly: visual orientation, algorithms, and identified neurons. *Journal of Neuroscience* 13:4563-4574.
- Egelhaaf M, Borst A, Reichardt W (1989) Computational structure of a biological motion-detection system as revealed by local detector analysis in the fly's nervous system. *Journal of the Optical Society of America A* 6:1070-1087.
- Fairhall AL, Lewen GD, Bialek W, de Ruyter van Steveninck R (2001) Efficiency and ambiguity in an adaptive neural code. *Nature* 41:787-792.

- Franceschini N, Riehle A, Le Nestour A (1989) Directionally selective motion detection by insect neurons. In: Stavenga DG and Hardie RC (eds) *Facets of Vision*, pp 360-390, Springer-Verlag, Berlin.
- Franz MO, Krapp HG (1998) Wide-field, motion-sensitive neurons and matched filters for optic flow fields. *Biological Cybernetics* 83:185-197.
- Götz KG (1964). Optomotorische Untersuchung des visuellen Systems einiger Augenmutanten der Fruchtfliege *Drosophila*. *Kybernetik* 2:77-92.
- Haag J, Egelhaaf M, Borst A (1992) Dendritic integration of motion information in visual interneurons of the blowfly. *Neuroscience Letters* 140:173-176.
- Harris RA, O'Carroll DC, Laughlin SB (2000) Contrast gain reduction in fly motion adaptation. *Neuron* 28:595-606.
- Harris RA, O'Carroll DC, Laughlin SB (1999) Adaptation and the temporal filter of fly motion detectors. *Vision Research* 39:2603-2613.
- Harrison RR, Koch C (2001) A silicon model of the fly's optomotor control system. *Neural Computation* 12:2291-2304.
- Hassenstein B, Reichardt W (1956) Systemtheoretische analyse der Zeit-, Reihenfolgen-, und Vorseichenauswertung bei der Bewegungspertzeption des Rüsselkäfers *Chlorophanus*. *Zeitschrift für Naturforschung* 11b:513-524.
- Hausen K (1993) The decoding of retinal image flow in insects. In: Miles FA and Wallman J (eds) *Visual Motion and its Role in the Stabilisation of Gaze*, Elsevier, London.
- Hausen K, Egelhaaf M (1989) Neural Mechanisms of Visual Course Control in Insects. In: Stavenga DG and Hardie RC (eds) *Facets of Vision*, pp 391-424, Springer-Verlag, Berlin.
- Hausen K (1982) Motion-sensitive interneurons in the optomotor system of the fly. II. The horizontal cells: Receptive field organization and response characteristics. *Biological Cybernetics* 46:67-79.
- Higgins CM, Douglass JK, Strausfeld NJ (2004) The computational basis of an identified neuronal circuit for elementary motion detection in dipterous insects. *Visual Neuroscience* 21:567-586.
- James AC (1992) Nonlinear operator network models of processing in the fly lamina. In: Nabet B (ed) *Nonlinear Vision*, pp 39-74, CRC Press, Boca Raton FL.
- Kern R, Lutterklas M, Petereit C, Lindemann JP, Egelhaaf M (2001) Neuronal processing of behaviourally generated optic flow: experiments and model simulations. *Network-Computation in Neural Systems* 12: 351-369.
- Kirschfeld K (1991) An optomotor control system with automatic compensation for contrast and texture. *Proceedings of the Royal Society of London B* 246:261-268.
- Kirschfeld K (1972) The visual system of *Musca*: studies on optics, structure, and function. In: Wehner R (ed) *Information processing in the visual system of arthropods*. Springer, Berlin Heidelberg New York, pp 61-74.
- Krapp HG, Hengstenberg B, Hengstenberg R (1998) Dendritic structure and receptive-field organization of optic flow processing interneurons in the fly. *Journal of Neurophysiology* 79:1902-1917.
- Laughlin SB, Weckström M (1993) Fast and slow photoreceptors – a comparative study of the functional diversity of coding and conductances in the diptera. *Journal of Comparative Physiology A* 172:593-609.
- Lipetz LE (1971) The relation of physiological and psychological aspects of sensory intensity. In: Loewenstein WR (ed) *Handbook of sensory physiology*. Springer, Berlin Heidelberg New York, pp 192-225.
- Maddess T, Laughlin SB (1985) Adaptation of the motion-sensitive neuron H1 is generated locally and governed by contrast frequency. *Proceedings of the Royal Society of London B* 225:251-275.
- Naka KI, Rushton WAH (1966) S-potentials from luminosity units in retina of fish (*Cyprinidae*). *Journal of Physiology (London)* 185:587-599.
- Poggio T, Reichardt W, Hausen K (1981) A neural circuitry for relative movement discrimination by the visual system of the fly. *Naturwissenschaften* 443:446.
- Potters M, Bialek W (1994) Statistical mechanics and visual signal processing. *Journal de Physique I* 4:1755-1775.
- Ruderman DL (1994) The statistics of natural images. *Network: Computation in Neural Systems* 5:517-548.
- Snyder AW (1979) Physics of vision in compound eyes. In: Autrum H (ed) *Comparative physiology and evolution of vision in invertebrates: Invertebrate photoreceptors*, vol. VII/6A of *Handbook of Sensory Physiology*. Springer-Verlag, Berlin, pp 225-313.

- Snyder AW, Stavenga DF, Laughlin SB (1977) Spatial information capacity of the eyes. *Journal of Comparative Physiology* 116:183-207.
- Srinivasan MV, Guy RG (1990) Spectral properties of movement perception in the dronefly *Eristalis*. *Journal of Comparative Physiology A* 166:287-295.
- Srinivasan MV, Laughlin SB, Dubs A (1982) Predictive coding: a fresh view of inhibition in the retina. *Proceedings Royal Society of London B* 216:427-459.
- Straw A, Rainsford T, O'Carroll D (2005) Estimates of natural scene velocity in fly motion detectors are contrast independent. *Current Biology*, submitted.
- Tolhurst DJ, Tadmor Y, Chao T (1992) Amplitude spectra of natural images. *Ophthalmology and Physiological Optics* 12:229-232.
- van Hateren JH, Snippe HP (2001) Information theoretical evaluation on parametric models of gain control in blowfly photoreceptor cells. *Vision Research* 41:1851-1865.
- van Hateren JH (1992) Theoretical predictions of spatiotemporal receptive fields of fly LMCs, and experimental validation. *Journal of Comparative Physiology A* 171:157-170.
- van Hateren JH (1997) Processing of natural time series of intensities by the visual system of the blowfly. *Vision Research* 37:3407-3416.
- van Hateren JH (1990) Directional tuning curves, elementary movement detectors, and the estimation of the direction of visual movement. *Vision Research* 30:603-614.

Appendix

To quantify the degree of saturation of a signal subjected to a limiting nonlinearity, we measured the fraction of time during the course of a simulation that the signal resides in the upper 10% or lower 10% of the range of the nonlinear function. The time spent in transition between these states is therefore consistent with the engineering concept of rise or fall time. We define moderate saturation as corresponding to 60% - 65% of simulated time spent in those parts of the range near the extrema, by the above criterion. This criterion was applied to both saturating nonlinearities, at early vision output and at EMD correlator output.

When a limiting nonlinearity was included in the signal path in simulations of the motion processing pathway, we measured and averaged the time spent near the extrema for the limited signal at five different latitudinal locations in the array, and adjusted the scaling of the signal at the input to the nonlinear block with the goal of obtaining moderate saturation. When motion adaptation was included in the model, a single scaling constant could be found that achieved moderate saturation for all five animated images in the test set, in all cases. However, when motion adaptation was not included, saturation necessarily fell outside the moderate range for some of the images in the set, due to the differing global contrasts. In these instances, the time spent near the extrema ranged from 48.5% to 52.3% for the lowest-contrast image *gardens*, and 82.3% to 89.0% for the image *hamlin*. The other three images were in or near the moderate saturation range.