

This paper was published in *Proceedings of SPIE* Vol. 5783, pp. 292–303, 2005 (SPIE Defense and Security Symposium, Infrared Technology and Applications Conference, Orlando, FL, March-April 2005), and is made available as an electronic reprint with permission of SPIE. One print or electronic copy may be made for personal use only. Systematic or multiple reproduction, distribution to multiple locations via electronic or other means, duplication of any material in this paper for a fee or for commercial purposes, or modification of the content of the paper are prohibited.

Insect-based visual motion detection with contrast adaptation

Patrick A. Shoemaker^a, David C. O'Carroll^b

^aTanner Research, Inc., 2650 East Foothill Blvd., Pasadena, CA 91107, USA

^bDiscipline of Physiology, University of Adelaide, Adelaide, SA 5005, Australia

ABSTRACT

The visual pathway that leads from the retina to the tangential cells in the third optical ganglion of the fly is a sophisticated system for the detection of visual motion. The tangential cells, whose responses are thought to characterize the state of egomotion of the animal, show a remarkable ability to encode velocity information about optic flow patterns to which they are sensitive, independent of the structure and contrast of viewed scenery. We describe a simulation study based on a model that accounts for key physiological features observed in the biological system, which contains nonlinear features that we expect to contribute to this capability. One of these features is motion adaptation, a phenomenon on which recent research has shed new light. We conclude that our models significantly reduce dependence of response on variable natural scenery, although they still do not perform as well in this respect as the biological neurons. This biological system has inspired an implementation of visual motion processing in analog VLSI technology. The neuromorphic circuits are intended for eventual on- or near-focal plane integration with photosensing. We describe the design approach and present results from preliminary versions of these circuits.

1. INTRODUCTION

Although insects are relatively simple organisms compared to vertebrates, with nervous systems of limited size and complexity, they nonetheless do a remarkable job of controlling flight and other behaviors based on their low-resolution visual sense. Detection of visual motion plays a critical role in these behaviors. A class of widely-studied cells in the third optical ganglion of the true flies (dipterans), the tangential neurons, displays sensitivity to patterns of optic flow across broad swaths of the visual field, and almost certainly comprises a system for estimation of the state of egomotion of the organism. Some of these cells, for example, have been linked to stabilization about the yaw axis in hovering flight¹, in certain species of flies. Other classes of neurons have also been found in this region of the brain that, although as yet less well-characterized, appear to be selective for different classes of visual motion; one in particular shows a remarkable selectivity for the motion of small targets against background².

We consider the chain of processing that leads from photoreception to the wide-field, egomotion-detecting tangential neurons in the insect brain. If this system is to provide the organism with information about its state of motion, it must estimate this information from time-varying patterns that depend on the absolute luminance, spectral content, contrast, and spatial structure of the visual scene, as transduced by the low-resolution imaging sensor that is the insect compound eye. The difficulty of this problem becomes clear when one considers that motion sensing should properly be *invariant* with respect to all of those parameters, because while they characterize the visual scene, they bear no relation to motion in or of that scene. They do, however, affect the response of neurons early in the visual pathway, and as well influence the response of models for the fundamental operations of motion detection in insects. In spite of this, tangential neurons display a remarkable insensitivity to variations of these parameters in natural scenery, which, although not entirely understood, seems due in large part to various stages of nonlinear processing in the signal path. We discuss this processing in brief, and focus in particular on the phenomenon of motion adaptation, which we have studied

experimentally (in the laboratory of D.O’C.) and modeled in simulations and in analog integrated circuitry (in the laboratory of P.S.).

2. VISUAL MOTION DETECTION IN INSECTS

Visual motion processing in insects begins with photoreception in the retina, and then proceeds through three successive optical ganglia, in order, the lamina, medulla, and lobula^{3,4}. The wide-field tangential neurons reside in a specialized portion of the lobula called the lobula plate.

2.1 Early vision

Vision begins in the insect compound eye, a hexagonal array of structures called ommatidia. Each ommatidium has an individual lenslet and a set of photoreceptor cells beneath. Compound eye optics are typically diffraction-limited, and they blur the image that appears on the insect retina to an extent that is well-matched to the spatial sampling period of the photoreceptors, preventing undersampling⁵. The photoreceptors adapt to the mean luminance in the environment, and give an approximately logarithmic response to changes in light intensity about their operating point⁶. Due to these characteristics, they can be thought of as signaling local *contrast* rather than absolute luminance. Fly photoreceptors are much faster than vertebrate receptors, with integration time constants on the order of a few milliseconds⁷. Bandpass temporal filtering occurs in the retina and lamina^{8,9}, along with some additional lowpass spatial filtering due to lateral interconnections in the lamina. Indirect results from motion-sensitive neurons further along in the visual pathway suggest that spatial highpass filtering may also occur in some species (i.e., the net spatial filter is bandpass in nature) (O’Carroll, unpublished observations). Insects generally possess color vision, but the pathways thought to be involved in motion detection are monochromatic¹⁰.

Collectively we refer to transduction and processing in the lamina as *early vision*.

2.2 Elementary motion detection

The specialized processing of visual motion is thought to begin in the medulla, which maintains a retinotopic architecture with columns of cells each associated with an overlying ommatidium. An early and still influential model of early motion detection, the correlational *elementary motion detector* (EMD), was formulated by Hassenstein and Reichardt in 1956¹¹. This model is based on a *correlation* of the signal associated with one ommatidium with the delayed signal from a neighboring ommatidium, as depicted schematically in Figure 1. The combination of the spatial connectivity, the delay operation, and the nonlinear interaction of the correlator endow the EMD with directional motion sensitivity. The final opponent stage depicted in Figure 1, which takes the difference of two mirror-image correlator outputs, enhances the directional properties and rejects the effects of temporal contrast not due to motion. The EMD produces a motion-related output without computing derivatives (a process that would amplify noise).

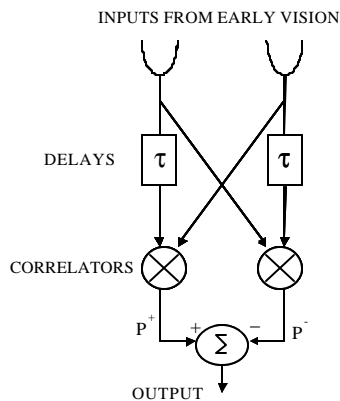


Figure 1: The Hassenstein-Reichardt or correlational elementary motion detector (EMD).

The correlational EMD model is well-supported by physiological and behavioral evidence, although its neural basis has been hard to pinpoint due to the technical difficulty of electrophysiological recording from the very small medulla cells that are believed to be involved. Evidence suggests that EMDs are formed between at least nearest neighbors and next-nearest neighbors on the hexagonal lattice, and are thus aligned with various directions in visual space¹². Physiological data in support of the EMD model have been taken primarily from more central neurons in the lobula complex, in particular the tangential cells themselves.

The output of a motion detector model of this kind 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*, however, the output 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 noise¹³.

2.3 Tangential cells: spatial integration

Tangential cells are large neurons with extensive, branched input structures, or dendrites, and they are believed to integrate signals from elementary motion detectors over wide areas of the visual field^{14,15,16}. These inputs may be weighted variably according to location and orientation of the EMD, with the result that the absolute sensitivity and directional preference of the tangential cell to visual motion can vary across its receptive field. In fact, this characteristic has led to the hypothesis that some tangential cells may act as *matched filters* for patterns of optic flow induced in the eye by particular modes of egomotion of the animal^{14,17}. In such cases, maps of directional sensitivity across the receptive field are similar to plots of optic flow vectors induced in the same area of the eye by the particular motion.

While tangential cells integrate input signals from a large number of EMDs, evidence suggests that this process is not linear. 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, cell response varies as the contrast or stimulus area is increased. However, for typical contrasts and full-field motion, this dependence disappears and the membrane potential saturates.

This characteristic has been modeled by Borst and colleagues¹⁸ as arising from synaptically-mediated ion conductances in the tangential cell membrane. They refer to this mechanism as ‘gain control’ (although it is a static nonlinearity and not gain control in the engineering sense). The details of the biophysics involved are beyond the scope of this paper, but the model may be analogized as a voltage division in which conductance to a positive (depolarizing) voltage source is mediated by ‘positive’ correlator outputs from EMDs (i.e., the signal P^+ in Figure 1), conductance to a negative (hyperpolarizing) voltage source is mediated by the ‘negative’ correlator outputs (i.e., P), and a fixed conductance is also connected to a neutral potential. Saturating behavior occurs when the variable conductances are large enough to dominate the fixed conductance.

2.4 Motion adaptation

When a tangential cell is exposed to strong motion stimuli, and subsequently probed with small (low motion-energy) test stimuli, the response to these test stimuli is reduced relative to that of an initially unstimulated neuron^{19,20,21}. This phenomenon has been termed motion adaptation. Recent work has produced strong evidence that this is due to a reduction in gain somewhere in the prior signal processing pathway, and also to a shift in the resting membrane potential of the cell when the adapting motion is in a direction that causes net excitation of the cell¹⁹. The gain reduction is elicited by motion in any direction, even when it causes little response, or net inhibition, in the cell. This gain reduction component has been shown to be spatially local (i.e., it occurs prior to integration of motion detector outputs) and to vary directly with contrast of the adapting stimulus, when that contrast is above some minimal threshold. Dependence on contrast, i.e., relative rather than absolute variations in luminance in the moving scene, is presumably due to the properties of the front-end photoreception.

The details of exactly what drives the process of motion adaptation remain a subject of research. We have modeled the gain reduction phenomenon with several approaches, including as a form of contrast gain control at the front end of the EMD operation. This simple approach has been the basis of an implementation of an adaptive EMD in analog VLSI circuitry. It is consistent with some of the basic features of the gain changes, although there are others that it does not explain: for example, adaptation seems to occur on multiple time scales, and it appears to be more strongly recruited by motion signals than by purely temporal contrast such as flicker.

2.5 Tangential cells: ‘velocity constancy’

Recent results from intracellular recordings in tangential cells have shown that their responses to a particular mode of visual motion (e.g., horizontal optic flow as would be induced by yaw rotation) are dependent on the speed of that motion, but are nearly invariant with respect to varying natural scenes. These results have been obtained from *in-vivo* experiments in which animals have been subject to moving stimuli while the intracellular recordings have been made. These results suggest that tangential cells respond to two characteristics of wide-field motion – the spatial pattern and the speed of optic flow – and that they are capable of rejecting the effects of differences in contrast and spatial structure in the natural environments in which the animal operates. This may be due in part to the fact that natural scenes tend to be relatively consistent in their spatial statistics^{22,23}. The spatial power spectra of luminance along one-dimensional paths in natural images follow an approximate $1/f$ characteristic, where f is spatial frequency^{24,25}. There is, however, still considerable variation in local structure and contrast from scene to scene. The correlational EMD that is used to model the initial stage of motion processing has a quadratic dependence on contrast of moving stimuli due to the presence of the correlation operation, and experiments confirm this second-order dependence in the biological system when contrast is very low¹⁵. Thus, velocity constancy must arise from other nonlinear properties of the system, and we assume that the nonlinear integrating properties of the tangential cells and motion adaptation are two contributing factors.

3. MODELING AND SIMULATION

3.1 A system-level model

In Figure 2 below is depicted an overview of the wide-field visual motion detection system, from photoreception to tangential cell, as modeled in this study. The lower part of the diagram refers to known nonlinearities in the processing chain. Motion adaptation has been putatively assigned a location between early vision and the elementary motion detection operations.

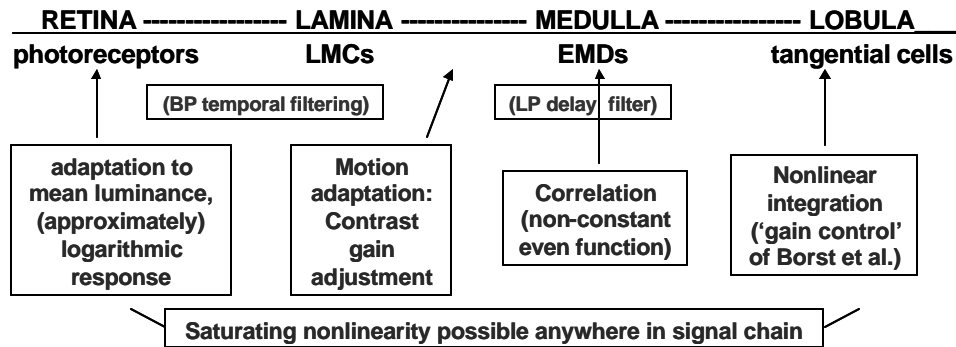


Figure 2: Overview of a system-level model of wide-field motion detection in the flies. At top are the brain structures involved in the processing and in the second line, neural or functional units. LMC here indicates the Lamina Monopolar Cell. Beneath them are indicated the temporal operators, and at the bottom are indicated postulated nonlinearities in this system.

The early vision model includes, in order, a first-order lowpass filter with corner frequency of about 20Hz; a Lipetz transformation, and a highpass filter with corner frequency of about 0.4Hz. The Lipetz transformation²⁶ is described by the equation

$$U = \frac{I^a}{I^a + I_0^a}, \quad (1)$$

where I is the input intensity, I_0 is a parameter defining mid-response level, a is an exponent between 0.5 and 1, and U is the output. We set the exponent to 0.7, and I_0 to 10, a realistic value for the numerical range of luminances in our input data (which were absolutely bounded by 0 and 255). The function in Equation (1) has been used to describe photoreceptor responses, and gives an approximately log-linear response over more than a decade of input intensity. No effort was made to model photoreceptor adaptation, given the limited dynamic range of our input data.

Saturating nonlinearities can presumably occur anywhere in the processing chain due to physical limits on the signaling ranges of the neural elements involved. Saturating nonlinearities have been included as elaborations in the correlational model for motion detection, and have been shown to improve reliability of velocity estimation²⁴. The Lipetz function in Equation (1) is inherently saturating; and we also inserted a saturating nonlinearity (a *tanh* function) in two locations, at the output of early vision, and at the outputs of the two EMD correlators. Gain was set so that ‘soft’ saturation occurred for typical input signals, i.e., the response characteristic was between linear and binary thresholding, with significant “squashing” for large signal excursions.

Motion adaptation was modeled as a form of gain control. Mean absolute deviation (MAD) of the early vision output, prior to the application of saturating nonlinearity, was estimated by full-wave rectification followed by a first-order lowpass filter with fixed (but user-selectable) time constant. The input signal was then divided by the MAD estimate.

In the elementary motion detector, the correlation operation was modeled as a pure multiplication, and the delay operator as a first-order lowpass filter with a time constant of 40ms. Using the phase delay of a lowpass filter is a typical approach, and is more consistent with the observed temporal tuning of the EMD than a pure time delay.

Tangential cell response was modeled as a function of the individual correlator outputs from a large array of EMDs. 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 the tangential cell. Then the wide-field neuron computation, with the Borst ‘gain control’ model for nonlinear integration, can be expressed in the form

$$V = \frac{\sum_{i,j} W_{ij} (P_{ij}^+ - P_{ij}^-)}{\sum_{i,j} W_{ij} (P_{ij}^+ + P_{ij}^-) + W_0}, \quad (2)$$

where V is the tangential cell output, W_{ij} are weights associated with each EMD, P_{ij}^+ and P_{ij}^- are the individual correlator outputs corresponding to P^+ and P^- in Figure 1, and the parameter W_0 , which is associated with a fixed membrane conductance, prevents singularity and limits variations in V due to noise when the denominator is small. W_0 was assumed negligible in our modeling since full-contrast, wide-field moving scenes were employed as stimuli.

It is by the particular choice of the weights in Equation (2) that a matched filter for a pattern of optic flow may be constructed. An optimal matched filter is obtained when each weight W_{ij} is set proportional to the projection of the expected velocity of optic flow at ommatidium j onto axis i . Here ‘expected velocity’ means the local velocity that is induced by the optic flow to be detected. In our case, we implemented a hexagonal array with EMDs formed between nearest neighbors only. Equal weights of unity value were assigned to EMDs aligned with the axes at 60° and 120° with respect to the vertical, across the entire receptive field. In this way, a ‘tangential cell’ with maximum sensitivity to uniform horizontal optic flow was constructed. In the biological system, this might represent a neuron with an equatorial receptive field that is sensitive to yaw rotation. Neurons with primary sensitivity to horizontal optic flow are found among the tangential cells in the flies, and are labeled HS cells¹.

3.2 Simulations of an ‘HS Cell’

We simulated the response of an artificial ‘HS cell’ to horizontal motion of natural scenery. The effective receptive field dimensions of the simulated array were about 40° in altitude by 160° in azimuth. Both the numerator in Equation (2), representing a straight sum of correlational EMD outputs, and the entire quotient, representing the HS cell output with nonlinear integration (Borst ‘gain control’), were computed on any given run. Two other key nonlinearities – saturation and motion adaptation – were enabled or disabled in various runs to evaluate their effects on model output. Finally, we also included spatial highpass filtering in some individual runs.

Five different, predominantly natural scenes from the habitat of several fly species were processed and animated to form the inputs for the simulations. These varied significantly in content and contrast. The 8bit green values were extracted from panoramic color photographs of the scenes (green being the part of the spectrum to which the photoreceptors involved in motion vision are most sensitive¹⁰). These data were blurred by convolution with a Gaussian kernel to approximate the spatial blurring imposed by the compound eye optics, and then resampled onto a hexagonal array with a spacing between centers of about 1.2°. This approximates the interommatidial angle in larger fly species such as hoverflies, which have relatively acute vision. The image was animated in uniform right-to-left motion over 720° of total ‘yaw rotation’, and time series of luminance values were extracted to serve as inputs to a time-domain simulation. The results reported here are for a single speed of motion, about 50°/s, which is near the velocity optimum or speed of maximal response of the EMD. The output of the model HS cell was averaged over one complete rotation for each run.

Simulations were performed using SPICE (Simulation Program with Integrated Circuit Emphasis), a tool for circuit simulations that also permits the implementation of abstract elements described by general equations, and which is optimized for the integration of stiff nonlinear differential equations.

3.3 Simulation Results

In Figure 3 below are shown average ‘HS cell’ outputs in response to each of the five animated images, for the basic model and for each of the test features individually. Each set of five data is normalized by its mean, so that the scatter can be compared directly from configuration to configuration. Because the mode of motion (horizontal translation) and speed are the same in all instances, this scatter is a measure of deviation from the ideal of velocity constancy, for any given model configuration.

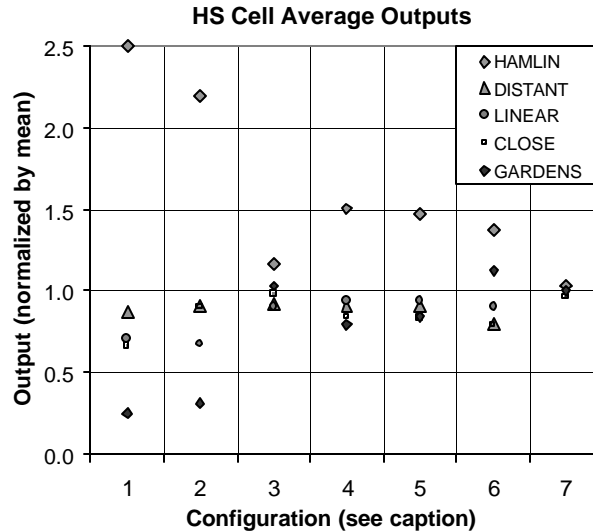


Figure 3: Variability of responses between different animated scenes of the model in various configurations, at the near-optimal speed of 50°/s. In each column of the graph are displayed responses to all five images for a particular model configuration, normalized by the mean over all five to allow direct comparison of the scatter for various model configurations. 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 of the images, including the ones with the highest (HAMLIN) and lowest (GARDENS) global contrasts, at a speed of 100°/s.

The data in the first column in Figure 3 illustrate the sensitivity of the basic correlational EMD to variations in image contrast: the mean ‘HS cell’ output varies by a factor of over ten between the highest and lowest contrast images. The three forms of nonlinear processing that we expect to reduce inter-scene variance – adaptation, saturation, and nonlinear integration in the HS cell – are all seen to do so to a significant extent. The final column (col. 7) in Figure 3 contains experimental data obtained from an actual HS cell in response to the same horizontal motion, for three of the five images in the data set (animated at 100°/s in this case). These images include those with the highest and lowest contrast, and it can be seen that there is less scatter in response to these in the biological neuron than with any of the model configurations tested.

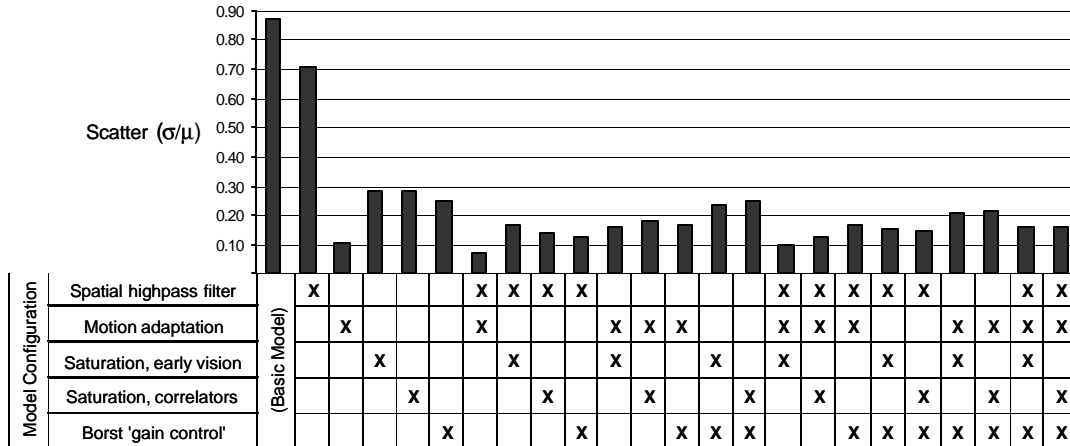


Figure 4: Summary of results from all simulations showing inter-scene scatter in model response to the five natural test images used in simulations. Headings at left indicate the test features, and an ‘X’ indicates the inclusion of a feature in the model configuration represented by each column. For 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.

In Figure 4 is shown a summary of all results from the simulation study. Of the various nonlinear features, the simple gain-control-based motion adaptation model consistently provides the greatest reduction in scatter; the four lowest values are achieved by configurations that include it. Interestingly, results obtained by varying the time constant of adaptation suggest that inter-scene variance is smallest for relatively fast adaptation, on the order of 100ms – 200ms for this speed of motion. This time scale is similar to the time constants in the temporal filters in the processing chain, and small enough so that there is significant interaction with individual signal excursions. This is very different from the usual situation with engineering gain control, in which gain changes are orders of magnitude slower than the frequencies contained in the signal being processed. When spatial highpass filtering is added to the model, it does not by itself greatly improve scatter, 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. Also of interest is the fact that 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 a fewer number of the features.

Finally, the results presented so far concern the global properties of the ‘HS cell’ with respect to velocity constancy as a performance standard. However, the local nature of gain changes in motion adaptation suggest that they may play a role in improving local representations of optic flow velocity. This would clearly allow improved performance in the context of matched filters for different patterns of optic flow, for example. The output of the correlational EMD in response to natural stimuli, in addition to being contrast dependent, is markedly unsteady over time, with a large variance relative to its mean. A central question is whether this variance can be reduced (relative to the mean) by the process of adaptation. We do not yet have quantitative data with regard to this hypothesis, but preliminary results suggest that even our simple gain control model for adaptation does result in increased uniformity of response across the receptive field in our test case of pure horizontal motion.

4. AN ANALOG VLSI MODEL FOR THE ADAPTIVE EMD

With recent interest in neuromorphic modeling of biological functions in analog VLSI technology, there have been several efforts to implement insect-inspired elementary motion detection models in analog integrated circuitry^{27,28,29}. These have been conceived as a means to achieve on- or near-focal plane processing of unsampled, continuous-time photosensor outputs to obtain motion information from moving imagery. Naturally, inclusion of the necessary processing on the focal plane places severe limits on the pixel density of a combined sensor/processor; however, it is important to note that the remarkable performance of flying insects is achieved with very low visual resolution: a compound eye which views almost an entire visual hemisphere may contain less than a thousand to a few tens of thousands of ommatidia, depending on the species. It is feasible with modern VLSI technology to consider integrated sensing / processing systems that view a portion of the visual field with resolution similar to that of some insects. In fact, an EMD-

based looming detector with an integrated array of only 17 by 17 photodetectors has been reported and applied to collision-avoidance purposes in a ground robot³⁰.

We have designed and built a series of analog integrated circuits that implement elementary motion detection with motion adaptation modeled on contrast gain control as discussed in the previous sections of this paper. To date, these chips have been configured with electrical inputs, but the designs are intended for eventual integration with bump-bonded or on-chip photodetectors. Both one-dimensional and hexagonal two-dimensional arrays of adaptive EMDs have been produced. In these arrays, the output currents of the individual EMD cells are summed together to form global outputs. The uniaxial EMD arrays are thus detectors of motion along their long axes, while in the two-dimensional array, a pair of detectors for orthogonal components of motion are formed by appropriate weighting and linear combination of EMD outputs. Higher-level processing in the form of the Borst ‘gain control’ model has also been fabricated but is not yet evaluated.

4.1 Adaptive EMD design

Our design includes circuits to approximate the processing of early vision, including temporal bandpass filtering and logarithmic transformation, gain control for adaptation, and the delay-and-correlate function of the EMD. These designs utilize MOS transistors biased in the weak inversion or subthreshold region of operation. Input currents are assumed to be 10nA or less, and the front-end circuitry is expected to operate over many decades of mean current levels below this maximum. Temporal lowpass filtering of the input is implemented by parasitic capacitance in the input circuitry, providing a response characteristic that varies with bias level (as do the responses of biological detectors). Temporal highpass filtering, as well as the temporal lowpass filters associated with adaptation and the delay operator in the EMD, are implemented with a current-mode, log-domain filter stage that is capable of achieving relatively long time constants³¹.

A copy of the current-mode signal that results from the initial bandpass filtering is full-wave rectified and lowpass-filtered to form an estimate of the mean absolute deviation (MAD) that is used for gain control in our model of adaptation. In a typical gain control circuit, linearity of the input-output relationship is preserved at any particular value of the gain. If I_{ac} represents an ac input signal, such as the output of the bandpass filter in our early vision circuitry, then an output is generated that is proportional to $I_{ac}/\|I_{ac}\|$, where $\|I_{ac}\|$ is a (positive-definite) estimate of amplitude (in this case, the MAD estimate). However, since we subject the vision signals to logarithmic compression, a different approach to signal normalization is adopted. A normalizing current proportional to $\|I_{ac}\|$ is added to the highpass-filtered signal, $I_{out} = I_{ac} + k*\|I_{ac}\|$, and this summed current drives an exponential load in the form of several diode-connected MOS transistors in series. This yields an output voltage proportional to $\log(I_{ac} + k*\|I_{ac}\|) = \log(\|I_{ac}\|) + \log(I_{ac}/\|I_{ac}\| + k)$, when $I_{ac} > -k*\|I_{ac}\|$. By proper choice of scaling constant k , signal excursions in which $I_{ac} = -k*\|I_{ac}\|$ can be made infrequent but not prevented, so a catch-diode is provided in the circuit to prevent large negative spikes in the output voltage. By driving a copy of the normalizing current $k*\|I_{ac}\|$ into an identical load, a differential output voltage may be generated that is proportional to $\log(I_{ac}/\|I_{ac}\| + k) - \log(k)$. This is a bipolar quantity that is zero when $I_{ac} = 0$, providing a gain-controlled, logarithmically compressed signal suited for the EMD operations to follow.

The differential output is applied to a pair of Gilbert multiplier circuits which compute the correlations that are part of the EMD function. It is also applied to a differential transconductance stage and converted to a current, which in turn is passed to a current-mode lowpass filter that produces the delayed version of the signal. Copies of this delayed signal are passed to Gilbert multipliers in neighboring cells, and the delayed signals from those neighbors are in turn passed back to the local cell. With proper inversion and biasing, these currents form the second set of inputs to each Gilbert multiplier in the EMD. The Gilbert multiplier and transconductance state naturally provide a saturating nonlinearity in their transfer characteristics, although the voltage signals are scaled so that these circuits are in their linear range for all but the highest contrast input segments.

4.2 Physical design approach

While analog neuromorphic VLSI design has been a subject of research since the late 1980’s, practically useful implementations in analog circuitry have remained scarce. This may be in part due to the predominance of digital signal processing technology, but it may also be due to the fact that nonidealities such as noise and transistor mismatches that present critical limitations for the performance of analog signal processing in general have not received a great deal of attention. Neuromorphic designs often exploit CMOS devices biased in the weak-inversion regime, in which sensitivity to variations in the built-in surface potential (‘threshold voltage mismatch’) is very high due to the exponential current-

voltage relation. Transistor-to-transistor variations in the built-in potential constitute the single most critical device nonideality affecting circuit performance in this regime³².

Device and layout-level approaches to reducing the effects of transistor mismatch include using large-area, physically adjacent transistors, and breaking up matched pairs into quads that can be arranged in common-centroid geometries. These approaches naturally increase circuit area. Maintaining similar physical/chemical neighborhoods around devices also promotes matching. Design approaches such as chopper or auto-zeroed amplifiers may also be used, but require even more area, and have additional disadvantages such as the introduction of clock signals that tend to contaminate subthreshold analog circuitry with switching noise.

We have taken a physical design approach of using large transistors laid out in regular arrangements in order to promote device matching. These measures, in addition to the signal averaging effect of summing large numbers of EMD outputs by tangential cell analogs, are intended to result in practically useful devices when motion detecting circuits are scaled up for practical applications. The silicon area consumed by the use of large devices (i.e., a standard channel size of $5\mu\text{m}$ square for matched transistors in a $0.35\mu\text{m}$ technology) is mitigated in our approach by routing and interconnecting nodes in metal layers over transistor gates, as in Figure 5.

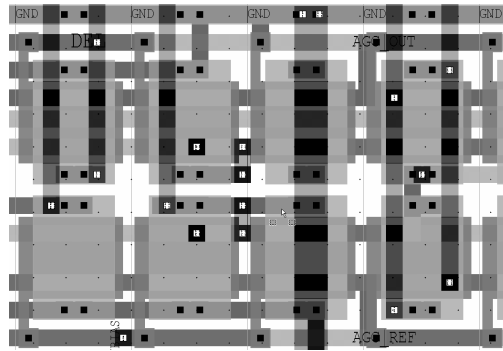


Figure 5: Segment of a layout, depicting routing of multiple layers of metal lines and interconnects over transistors, which are the rectangular features in background.

Regularity of layout is illustrated by a template cell for a hexagonal array as depicted below in Figure 6. In this layout, every transistor in a vertical column has an identical (or near-identical) physical/chemical neighborhood:

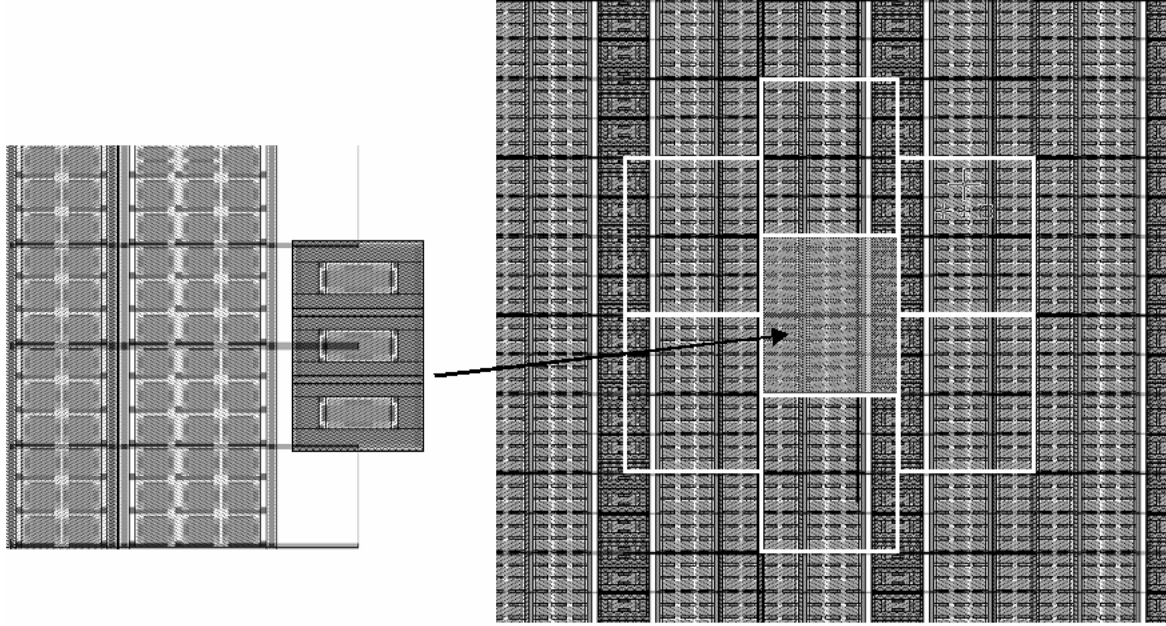


Figure 6: Template for a cell (at left) suitable for layout in a regular hexagonal array (at right). The vertical elements of the template consist of, from left to right, PMOS transistors (in N-well), NMOS transistors, and capacitors and transistors in an independent N-well.

4.3 Experimental results

As a fundamental test of the effectiveness of adaptation in the EMD, we have applied input sequences corresponding to moving sinusoidal gratings to EMD arrays on the chips. The contrast and speed of motion of these sinusoidal gratings were varied to evaluate the velocity and contrast dependence of the circuits. Because the spatial statistics of the gratings are stationary during any test, the steady-state response of the circuitry may be used to directly evaluate its compliance with velocity constancy.

We performed such tests using a computer-controlled multi-channel digital-to-analog conversion board to supply inputs to the device under test. The voltage amplitude of the channel outputs and the effective speed of motion of the grating could be controlled by the software driving the DAC hardware, using entered parameters. The DAC output voltages were lowpass-filtered to remove quantization noise, using passive components on the test board carrying the EMD chip. The voltages were converted to currents by on-chip transconductance amplifiers, in order to mimic the current-mode inputs that would be supplied by photodiodes operating in continuous time. By taking the output of a single arm of a differential stage, the zero-mean voltage inputs were converted to single-ended (positive) currents. A separate transistor driven by one of the input channels was used to provide a copy of an input current that could be monitored to evaluate the inputs to the EMD circuits themselves. The transistors were biased such that the mean input current level was $4\text{nA} - 5\text{nA}$.

Strictly speaking, this scheme does not result in input signals to the EMD whose statistics (other than amplitude) are stationary from run to run, since the harmonic content of these inputs depend upon input signal amplitude due to the nonlinearity imposed by the logarithmic transformation in early vision. Nevertheless, the same circumstances prevail in the natural system.

Some results are depicted below in Figure 7 below, for a uniaxial EMD array. The overall array is composed of two linear 8×1 subarrays with 7 equivalent EMDs in each. The two subarrays are concatenated together for reduction of noise and random errors. They are independent except that inputs to the two units in any column are common, and the outputs of all EMDs are summed together. Motion was in the row-wise direction, and the sine wave stimulus had a spatial period of $\frac{1}{4}$ cycle per pixel, which is the spatial optimum for the EMD. The EMD temporal filters were tuned to corner frequencies 3 – 4 times higher than those in the insect visual system, for ease of testing.

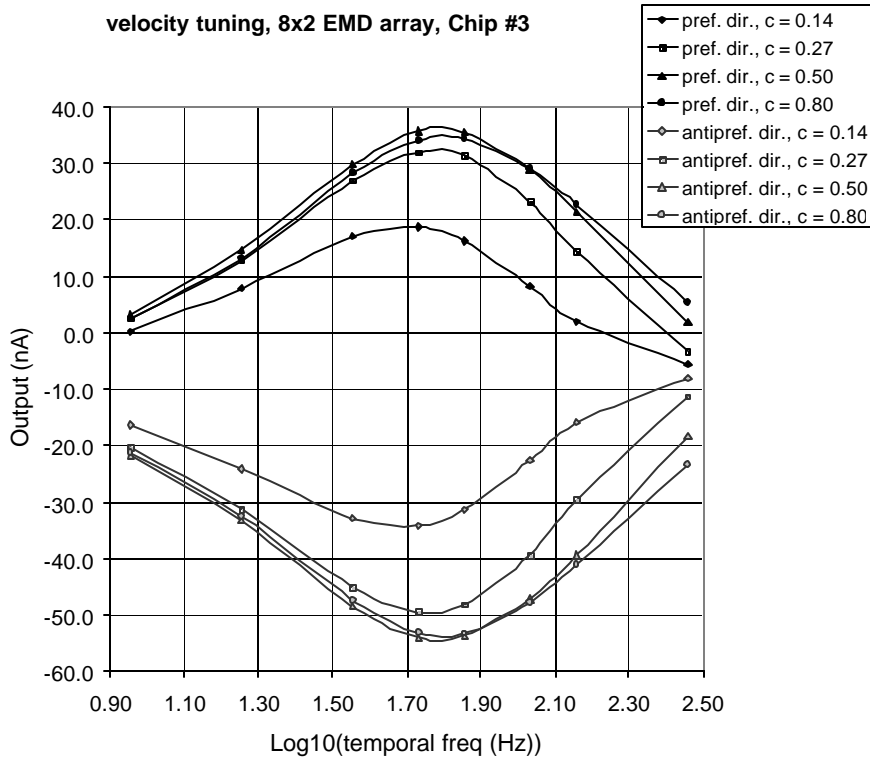


Figure 7: Velocity tuning curves for a uniaxial EMD array, in response to a moving sine wave grating stimulus of period $\frac{1}{4}$ cycle per pixel. Responses to motion in each direction ('preferred' and 'antipreferred') are depicted. The parameter c is the contrast of the EMD input, calculated as the RMS amplitude divided by the mean value of the current supplied to the EMD circuits. The temporal frequency of the stimulus, which is proportional to its velocity, is shown in log scale on the abscissa. The ordinate represents mean output. The EMD delay filter was tuned to maximize output at a temporal frequency of 70Hz, while the time constants for the temporal highpass filter in early vision and the lowpass filter in the contrast gain control were both set to approximately 50ms.

The curves in Figure 7 show the typical, unimodal shape that is seen in the HS cell response as a function of image velocity. For higher contrast values, the curves nearly overlap (particularly in the range of velocities below the optimum value at which the peak response occurs), reflecting the effects of adaptation. The results of circuit nonidealities are seen in Figure 7 in the form of an offset error of about -8nA , and in the gain control circuits falling out of compliance at lower contrasts. This begins for contrast values in the range of 0.30, and is strongly evident for the lowest contrast stimuli at $c = 0.14$. Typical contrast for natural scenery is on the order of 0.40 – 0.50, and for the natural scenes used in our simulation study, the contrast computed by the RMS method ranged from 0.34 – 1.18. The circuit nonidealities evident in Figure 7 can both be traced to mismatch in the built-in surface potential between matched pairs (or quads) of transistors.

5. CONCLUSIONS

We have considered a system in which spatiotemporal patterns of imagery are analyzed to obtain information about visual motion: that is, the visual motion processing pathway that leads from photoreception to the wide-field tangential cells in the third optical ganglion of the fly. These cells are a class of neurons that are thought to encode a representation of the state of egomotion of a flying insect. We have developed a model based on this system and examined its response to inputs corresponding to animated natural imagery. Our approach accounts for the basic nonlinear and temporal bandpass characteristics of early vision, includes a simple model of motion adaptation as a form of contrast gain control, followed by a correlational elementary motion detector, and finally includes a model of the tangential cell itself that performs a nonlinear integration of the outputs of an array of EMDs over its receptive field in visual space. A saturating nonlinearity is included prior to the EMD to reflect the limited ranges of the physical signals in the system. Three of the critical nonlinearities in this model – saturation, motion adaptation, and nonlinear integration – we expect contribute to

the remarkable capability of biological tangential cells to reject differences in the contrast and structure of natural scenes, and reliably signal the speed of a particular mode of motion. We examined the response of our model to uniform horizontal translation of five different natural images, and conclude that each of the critical nonlinearities significantly reduces the inter-scene variance of the model response, but not to the degree seen in the biological neuron, and that combinations of these nonlinearities do not further improve performance. Motion adaptation with relatively short time constants – similar in time scale to the other temporal filters in the system and the signals themselves – was found to give the greatest reduction in inter-scene variance. We speculate that unmodeled characteristics of motion adaptation are likely to contribute to the difference in performance between the model and biological neurons. Finally, preliminary results suggest that the spatially local nature of adaptation may increase the quality of information about local optic flow in the visual field.

This biological system has served as the inspiration for the ongoing development of a biomimetic motion detection system in analog VLSI CMOS circuitry. We have demonstrated uniaxial arrays of adaptive EMDs with response characteristics similar to those of biological tangential neurons. The inclusion of the contrast gain control model for motion adaptation allows these circuits to respond to variations in speed of motion while largely rejecting variations in contrast of the stimuli. Our design approach involves regular layout of large-gate-area transistors in order to minimize the effects of mismatches in transistor characteristics on circuit performance in the weak-inversion regime in which the circuits are designed to operate. Larger arrays will be integrated with photosensing in future efforts.

REFERENCES

1. Hausen, K. & Egelhaaf, M. "Neural Mechanisms of Visual Course Control in Insects," in *Facets of Vision*, Stavenga, D. & Hardie, R. (eds.), pp. 360-390, Springer-Verlag, Berlin, 1989.
2. O'Carroll, D. "Feature-detecting neurons in dragonflies," *Nature* **362**, pp. 541-543, 1993.
3. Clifford, C.W.G. & Ibbotson, M.R. "Fundamental mechanisms of visual motion detection: models, cells and functions," *Progress in Neurobiology* **68**, pp. 409-437, 2003.
4. Egelhaaf, M. & Borst, A. "A look into the cockpit of the fly: visual orientation, algorithms, and identified neurons," *Journal of Neuroscience* **13**, pp. 4563-4574, 1993.
5. Snyder, A.W. "Physics of vision in compound eyes," in *Comparative physiology and evolution of vision in invertebrates: Invertebrate photoreceptors*, in *Handbook of Sensory Physiology*, Autrum, H. (ed.), vol. VII/6A, pp. 225-313, Springer-Verlag, Berlin, 1979.
6. van Hateren, J.H. "Processing of natural time series of intensities by the visual system of the blowfly," *Vision Research* **37**, pp. 3407-3416, 1997.
7. Laughlin, S.B. & Weckström, M. "Fast and slow photoreceptors – a comparative study of the functional diversity of coding and conductances in the diptera," *Journal of Comparative Physiology A* **172**, pp. 593-609, 1993.
8. James, A.C. "Nonlinear operator network models of processing in the fly lamina," in *Nonlinear Vision*, Nabet, B. (ed.), pp. 39-74, CRC Press, Boca Raton FL, 1992.
9. Srinivasan, M.V., Laughlin, S.B., & Dubs, A. "Predictive coding: a fresh view of inhibition in the retina," *Proceedings Royal Society of London B* **216**, pp. 427-459, 1982.
10. Srinivasan, M.V., Guy, R.G. "Spectral properties of movement perception in the dronefly *Eristalis*," *Journal of Comparative Physiology A* **166**, pp. 287-295, 1990.
11. Hassenstein, B., Reichardt, W. "Systemtheoretische analyse der Zeit-, Reihenfolgen-, und Vorseichenauswertung bei der Bewegungserzeption des Rüsselkäfers *Chlorophanus*," *Zeitschrift für Naturforschung* **11b**, pp. 513-524, 1956.
12. Buchner, E. "Elementary movement detectors in an insect visual system," *Biological Cybernetics* **24**, pp. 85-101, 1976.
13. Potters, M., Bialek, W. "Statistical mechanics and visual signal processing," *Journal de Physique I* **4**, pp. 1755-1775, 1994.
14. Krapp, H.G., Hengstenberg, B., Hengstenberg, R. "Dendritic structure and receptive-field organization of optic flow processing interneurons in the fly," *Journal of Neurophysiology* **79**, pp. 1902-1917, 1998.
15. Egelhaaf, M., Borst, A., Reichardt, W. "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**, pp. 1070-1087, 1989.
16. Franceschini, N., Riehle, A., Le Nestour, A. "Directionally selective motion detection by insect neurons," in *Facets of Vision*, Stavenga, D.G. and Hardie, R.C. (eds), pp. 360-390, Springer-Verlag, Berlin, 1989.
17. Franz, M.O., Krapp, H.G. "Wide-field, motion-sensitive neurons and matched filters for optic flow fields," *Biological Cybernetics* **83**, pp. 185-197, 1998.
18. Borst, A., Egelhaaf, M., Haag, J. "Mechanisms of dendritic integration underlying gain control in fly motion-sensitive neurons," *Journal of Computational Neuroscience* **2**, pp. 5-18, 1995.
19. Harris, R.A., O'Carroll, D.C., Laughlin, S.B. "Contrast gain reduction in fly motion adaptation," *Neuron* **28**, pp. 595-606, 2000.
20. de Ruyter van Steveninck, R., Zaagman, W.H., Mastebroek, H.A.K. "Adaptation of transient responses of a movement-sensitive neuron in the visual system of the blowfly *Calliphora erythrocephala*," *Biological Cybernetics* **54**, pp. 223-236, 1986.

21. Maddess, T., Laughlin, S.B. "Adaptation of the motion- sensitive neuron H1 is generated locally and governed by contrast frequency," *Proceedings of the Royal Society of London B* **225**, pp. 251–275, 1985.
22. Ruderman, D.L. "The statistics of natural images," *Network: Computation in Neural Systems* **5**, pp. 517-548, 1994.
23. Tolhurst, D.J., Tadmor, Y., Chao, T. "Amplitude spectra of natural images," *Ophthalmology and Physiological Optics* **12**, pp. 229-232, 1992.
24. Dror, R.O., O'Carroll, D.C., Laughlin, S.B. "Accuracy of velocity estimation by Reichardt correlators," *Journal of the Optical Society of America A* **18** pp. 241-252, 2001.
25. van Hateren, J.H. "Processing of natural time series of intensities by the visual system of the blowfly," *Vision Research* **37**, pp. 3407-3416, 1997.
26. Lipetz, L.E. "The relation of physiological and psychological aspects of sensory intensity," in *Handbook of Sensory Physiology*, Loewenstein, W.R. (ed.), Springer, Berlin Heidelberg New York, pp 192–225, 1971.
27. Harrison, R.R., and Koch, C. "A robust analog VLSI Reichardt motion sensor," *Analog Integrated Circuits and Signal Processing* **24**, pp. 213-229, 2000.
28. Delbrück, T. "Silicon retina with correlation-based velocity-tuned pixels," *IEEE Transactions on Neural Networks* **4**, pp. 529-541, 1993.
29. Sarpeshkar, R., Bair, W., and Koch, C. "An analog VLSI chip for local velocity estimation based on Reichardt's motion algorithm," in *Advances in Neural Information Processing Systems*, Hanson, S., Cowan, J., and Giles, L. (eds.), pp. 781–788, Morgan Kaufman San Mateo, CA, vol. 5, 1993.
30. Harrison, R. "A low-power analog VLSI collision detector," in *Advances in Neural Information Processing Systems*, Thrun, S., Saul, L., Schölkopf, B. (eds.), MIT Press, Cambridge, vol. 16, 2003.
31. Shoemaker, P.A. "A methodology for long time constant log-domain filters in CMOS," *Analog Integrated Circuits and Signal Processing*, in press, 2005.
32. Pavasovic, A., Andreou, A.G., Westgate, C.R., "Characterization of subthreshold MOS mismatch in transistors for VLSI systems," *Journal of Analog Integrated Circuits and Signal Processing* **6**, pp. 75-85, 1994.