Elaborated Reichardt correlator for velocity estimation tasks

Sreeja Rajesh\textsuperscript{a,b,c}, David O’Carroll\textsuperscript{a,b,c} and Derek Abbott\textsuperscript{b,c}

\textsuperscript{a}Department of Physiology and
\textsuperscript{b}Department of Electrical & Electronic Engineering and
\textsuperscript{c}Centre for Biomedical Engineering, The University of Adelaide, SA 5005, Australia

ABSTRACT

The study of insect vision is believed to provide a key solution to many different aspects of motion detection and velocity estimation. The main reason for this is that motion detection in the fly is extremely fast, with computations requiring only a few milliseconds. So the insect visual system serves as the basis for many models of motion detection. The earliest and the most prominent model is the Reichardt correlator model. But it is found that in the absence of additional system components, the response of a simple Reichardt correlator model is dependent on contrast and spatial frequency. Dror has demonstrated that the addition of spatial and temporal filtering, saturation, integration and adaptation in a correlator based system can improve its performance as a reliable velocity estimator.

In this paper, we try to further investigate and expand his model to improve the correlator performance. Our recent neurobiological experiments suggest that adaptive mechanisms decrease EMD (elementary motion detector) dependence on pattern contrast and improve reliability. So appropriate modelling of an adaptive feedback mechanism is done to normalise contrast of input signals.

1. INTRODUCTION

Flying insects are capable of navigating through a complex environment with remarkable speed and accuracy, despite their small size and the relative simplicity of their brains. A number of their important navigational tasks rely mainly on visual information. The insect visual system is an excellent example of a robust motion system that works in a natural environment. Flies have a remarkable ability to navigate and adapt quickly to an unstructured environment through the motion information that they get from their low resolution compound eyes. The study of biological neural systems alone provides numerous examples of computational systems that are far more complex than any man made system and perform real time sensory and motor tasks in a manner that humbles the most advanced artificial systems.\textsuperscript{1} The neural pathway of the fly visual system is the best understood neural pathway in any animal. So a number of biologically inspired algorithms for motion detection have been proposed in biological and computer vision literatures.\textsuperscript{2–6}

The earliest and probably the most famous model of motion detection inspired by biological systems was developed by Reichardt and Hassentein\textsuperscript{7} in 1956 after a series of behavioural experiments examining the optomotor response of insects. The Reichardt or correlation motion detector has a highly parallel architecture. Each elementary motion detector (EMD) detects motion in a preferred direction by comparing a signal from one receptor with a delayed signal from the other receptor. The comparison is performed using a nonlinear, multiplicative interaction between the two channels. Two mirror symmetric EMDs tuned to opposite directions are subtracted to form a bidirectional motion detector. The multiplicative interaction employed in this detector is an excitatory mechanism.

Several authors have proposed elaborations of the ‘basic’ Reichardt detector by either adding temporal, spatial or spatio-temporal filters before applying the input to the EMD cell.\textsuperscript{8,9} In one of the most recent elaborations of this model, Dror,\textsuperscript{10} suggests that the inclusion of additional system components to perform pre-filtering,
response compression, integration and adaptation to a basic Reichardt correlator can make it less sensitive to contrast and spatial structure thereby providing a more robust estimate of local image velocity. It was found from the data obtained via intracellular recordings of the steady-state responses of wide field neurons in the hoverfly Volucella, that the shape of the curves obtained agreed well with the theoretical predictions made by Dror. But Dror’s model remains contrast dependent.

In this paper, an attempt is made to reduce the dependence of the Reichardt correlator model to contrast so that it can give more accurate velocity information. Firstly, the correlational EMD model for the early stage of motion detection in insects, is elaborated to mimic the properties of the fly visual system. Then based on motion adaptation studies in the insect visual system,\textsuperscript{11,12} contrast gain reduction is implemented using a feedback mechanism. It is found that, the resulting adaptive feedback EMD array model is successful in making the response less sensitive to contrast.

2. THE INSECT VISUAL SYSTEM

Compared to other animals, insects have a visual system of intermediate complexity. The secret of the simplicity of insect’s visual system is that unlike other animals like human beings where a complete image is transferred to the brain and a huge neural structure extracts what is perceived as an image, in insects, however, many of the tasks, like the perfect recognition of objects, have either been eliminated or simplified. So in the visual system of insects, motion detection plays a predominant role. The insect visual system is classified into four layers, retina, lamina, medulla and lobula as shown in Fig. 1. Each of these ganglia in the optic lobe are organised into columns and strata.

![Cross section of the fly's brain with the compound eyes. From Egelhaaf and Borst(1992).](http://proceedings.spiedigitallibrary.org/)

**Fig. 1.** Cross section of the fly's brain with the compound eyes. From Egelhaaf and Borst(1992).

2.1. Retina

The retina has a large array of photoreceptors which sample the visual field. Each of a fly’s two compound eyes consists of several thousand facets, or ommatidia, arranged in a vertically oriented hexagonal lattice.\textsuperscript{13} The ommatidium is a functional unit comprising a dioptric apparatus and a rhabdom, which accommodates eight photoreceptors. The light captured by each ommatidium is projected onto the eight photoreceptors. Each compound eye of the insect is composed of approximately 3000 to 4000 ommatidia. The ommatidial sampling lattice forms the basis for motion detection. Behavioural experiments indicate that stimulation of two individual photoreceptors in nearby ommatidia is sufficient to produce a turning response.\textsuperscript{14} These results were confirmed by
Riehle and Franceschini\textsuperscript{15} in neural recordings from motion sensitive neurons, which suggests that the two arms of the correlator corresponds to nearby ommatidia, with $\Delta \phi$ depending on the inter-ommatidial angle. Studies on phototransduction in insects reveal that the retinal signal, as sampled by the photoreceptors, is already blurred by diffraction effects of the lens optics as well as the properties of the photoreceptor themselves.\textsuperscript{16} Hence the response of the photoreceptor to an input image has the characteristics of spatial low pass filter. Incidentally this spatial filter is matched to the sample ‘mosaic’, for optimal aliasing. Also it is found that some temporal pre-filtering takes place in the photoreceptors. Photoreceptors, which provide the inputs to the correlator, depend on chemical transduction processes\textsuperscript{17} that cannot respond instantly to changes in the luminance signal. They therefore filter their inputs with a temporal impulse response found experimentally to have a log-normal form lasting several tens of milliseconds.\textsuperscript{18, 19}

2.2. Lamina

The first visual ganglion, the lamina, has a large number of identical channels. The main output cells of the lamina, the monopolar cells, codes contrast reduce redundancy due to large variation in background intensity. Contrast enhancement is thus a vital function of the lamina and this is performed through biological spatial high pass filters (lateral inhibition). Anatomical evidence suggests that the large monopolar cells (LMCs) form the primary input channel to the correlators.\textsuperscript{20} In addition, there is clear evidence that some kind of temporal highpass filtering takes place in the lamina.\textsuperscript{17, 21} Recordings and modelling by James et al from the temporal impulse responses of LMCs (in the hoverfly \textit{Eristalis}), reveals that the LMCs have the characteristic of the difference of two log normal functions.\textsuperscript{21} While some LMCs also appear to perform high pass spatial filtering, the phenomena is weaker and more variable than high pass temporal filtering.\textsuperscript{17, 21}

2.3. Medulla

The medulla is one of the most complex layers of the insect visual system, characterized by numerous classes of local, columnar neurons as well as ‘tangential cells’, which arborize across large number of medulla columns. Because of their small size, little is known about the physiology of most medulla neurons. Most of our understanding of properties of such cells comes from studying their target neurons in the lobula and lobula plate, which are large enough for physiological recordings. Nevertheless, some evidence suggests that columnar neurons in the medulla are responsible for key stages of local motion detection. Douglass and Strausfeld\textsuperscript{20} recorded a number of small field retinotopic neurons in the medulla, lobula and lobula plate. They found several motion sensitive cells whose physiological and anatomical characteristics suggest that they serve as components of an array of EMDs.

2.4. Lobula

The lobula receives input from the medulla to perform higher level functions. The lobula contains wide field directionally selective motion detection neurons which are responsible for the detection of whole field or whole frame motion. In the lobula, the output of the local retinotopic movement detectors is pooled. This is accomplished in the dendrites of the tangential cells in the lobula plate. Intracellular recordings from the wide field cells indicate that they sum the outputs of local Reichardt correlators in their receptive fields.\textsuperscript{22}

3. REICHARDT CORRELATOR MODEL

Fig. 2 shows a simplified version of the Hassenstein and Reichardt correlator model. Both behavioural and physiological experiments on the insect visual system indicate that movement detection occurs between neighbouring points of the sampling lattice of the eye. This implies that a motion detector has two input channels and the two channels should possess an asymmetric arrangement. Asymmetry is necessary for the detector to acquire direction selectivity. Furthermore the interaction between two channels must be nonlinear if the detector is to respond selectively to moving gratings. The asymmetrically nonlinear interaction between two input channels is termed the elementary motion detector (EMD).\textsuperscript{23} Such an EMD will elicit a strong response when a visual stimulus moves in a specified direction (the preferred direction) and weak response when the stimulus moves in the opposite direction.

A basic Reichardt correlator is formed by combining two EMDs that are tuned to opposite direction, to indicate bidirectional motion. The nonlinear interaction is assumed to be a multiplication, $M$, which is the
Fig. 2. The Reichardt correlator has two arms A and B. Each of the two time dependent inputs \( s_1 \) and \( s_2 \) has a fixed angular separation \( \Delta \phi \), and pass through a linear delay filter \( D \) before being multiplied by the other, undelayed signal. The results of the two correlations thus obtained ‘cor A’ and ‘cor B’ are subtracted to produce a single output ‘cor’. An object moving to the right will produce a positive output; an object moving to the left will produce a negative output.

The simplest possible nonlinear interaction. The asymmetry is implemented by using a delayed unit, \( D \), which is a first order low pass filter. This filter acts as a delayed element by introducing phase shift of up to 90°. The input channels are spatially separated by an angular distance, \( \Delta \phi \). The delay filter \( D \) determines the temporal frequency tuning and therefore the response to different motion velocities at any given spatial frequency. The outputs of the two multiplications are subtracted to give a single time-dependent correlator output.

Though insects and humans appear to be capable of estimating image velocities, the basic correlator model does not function as a velocity estimator. It reliably indicates directional motion of sinusoidal gratings, but the response depends on contrast (brightness) and spatial frequency (shape) as well as velocity. Furthermore, since the EMD is a complex spatiotemporal bandpass filter, any one pattern produces same response at two unique velocities.

4. NATURAL STIMULI

Natural images are not arbitrary. Recent work has shown that certain image statistics are highly predictable in the natural world and that the biological visual system is optimized to take advantage of these statistics. Although the simple correlator model produces more meaningful estimates of velocity for natural images than for arbitrary sinusoids, it suffers from two major shortcomings. First, the standard deviation of the correlator output is huge relative to its mean (3.3 to 76 times). Second, the mean correlator response for most natural images peaks at an angular velocity of 35 – 40°/s. Because the velocity range below the peak response corresponds to the most probable range of inputs, in the absence of contradictory information a correlator response is interpreted as the lower of the two putative velocities. Image velocities above the peak will therefore be misinterpreted. A shorter delay filter time constant would raise the peak response velocity, but experimentally described time constants are not sufficiently low to account for the fact that insects may turn and track targets at velocities up to hundreds of degrees per second. It is found that additional physiological components (temporal high pass filtering and spatial low pass filtering) help to overcome these problems, raising the peak response velocity.
Fig. 3. The elaborated version of the fly EMD model with spatial and temporal pre-filtering. Spatial filtering is done using a Gaussian filter of half width (standard deviation) 2° and temporal filtering is done with the difference of log normal filters. Motion is detected locally by correlation method and is pooled so as to enable wide field motion detection.

Fig. 4. The elaborated EMD array. In our EMD array model, an array of 240 elaborated EMDs are used to detect motion. The input stimulus given is an image of width 8352 pixels, and considering the inter ommatidial angle as 1.5° there will be typically 240 ommatidia (EMDs) working together to detect motion. The output of these EMDs is pooled to enable wide field motion detection.
and lowering the relative error of the correlator output. Analysis and simulations suggest that the processes
commonly found in visual systems, such as pre-filtering, response compression, integration, and adaptation,
improve the reliability of velocity estimation and expand the range of velocities coded.\textsuperscript{11, 13, 17, 30–33}

Following on from this work, we propose here an additional elaboration of this model to take account of recent advances in our understanding of non-linear adaptive properties of insect motion detection.\textsuperscript{11, 12}

**Fig. 5.** The panoramic natural image given as stimulus to the EMD model. A panorama of the image is formed by ‘warping’ 12 image tiles at 30\degree intervals to remove lens distortions and then by wrapping its ends together using Apple Quicktime VR software on a Macintosh computer.

**5. ELABORATED MOTION DETECTOR MODEL**

In our experiments, natural images photographed from favoured hovering positions of the hoverfly are used. Then the edges of the images are wrapped to form a panoramic image which is given as a stimulus to our fly EMD model. A panorama is formed by ‘warping’ 12 image ‘tiles’ at 30\degree intervals to remove lens distribution and then wrapping its ends together. This is done using Apple Quicktime VR software on a Macintosh computer. The resulting image has a width of 8352 pixels and height of 1264 pixels. Spatial pre-filtering is implemented by two-dimensional convolution of the image with a Gaussian kernel of half width (standard deviation) of 2\degree, which approximates the acceptance function of typical fly photoreceptor.\textsuperscript{34} Only the luminance (gray scale) information is taken from the image using the green channel since photoreceptors are green sensitive. The spatially low pass filtered image is illustrated in Fig. 6. The distance between two ommatidia in an insect eye is between 1 and 1.5 degrees. Since the insect is looking at an image of width 8352 pixels in 360 degrees, if we consider the inter-ommatidial angle as 1.5 degrees, there will be an array of typically 240 ommatidia (EMDs) working together to detect motion as shown in Fig. 4.

**Fig. 6.** The natural image is pre-filtered using a spatial low pass filter, imitating the characteristics of the photoreceptor. The spatial filter used here is a Gaussian filter of half width (standard deviation) two degrees.

This image is temporally filtered with a difference of log-normal filter to copy the response of the lamina monopolar cells.\textsuperscript{18, 19, 21}

The temporally pre-filtered image is then converted to a space-time matrix based on desired velocity, which is then simulated along the height of the image in pixels. The height of the image is also divided into 36 ommatidia by calculating the pixels per degree and keeping the inter-ommatidial angle as 1.5 degrees.
Fig. 7. shows a single row of the image animated at a constant velocity of $200^\circ/s$ and sampled onto an array of 240 ommatidia.

![Row of 240 Ommatidia](image)

Fig. 8. The input and the output responses of a single EMD from the modelled array moving with constant velocity of $200^\circ/s$. Refer to Fig. 2 for explanation of each component. Note that the EMD output as a function of time is highly variable, and is maximal whenever high contrast features more across the EMD inputs ($s_1, s_2$).

![Response](image)

Then this spatio-temporally pre-filtered image is given to the EMD array which correlates the inputs to give an array of outputs as done in the insect eye. Then the EMD array model copies the lobula by averaging the outputs to produce an average EMD response. The simulations are first performed with a constant velocity of 200 degrees per second and the input output responses are shown in Fig. 8. Then the simulations are repeated by increasing the velocity in steps and the mean correlator response is shown in Fig. 9. It is found that at higher velocities, the response decreases a little.

6. MOTION ADAPTATION IN THE FLY VISUAL SYSTEM

When you observe a fly hovering near a flower, you will be struck by the ability of its visual system to estimate self motion in order to achieve stabilization during hovering, which shows its sensitivity to low velocities. Similarly when you watch these insects engaging in high speed aerial pursuits while chasing mates, it is clear that their
visual system adapts easily to higher velocities as well. These activities of insects demand a visual system with a large dynamic range. Although physiological recordings demonstrate that insect motion detectors have such high sensitivity to contrast that they are able to respond over a huge range of velocities, this inherently high sensitivity to motion makes them prone to saturation.\textsuperscript{11,12}

Several electrophysiological investigations into the dynamic response properties of fly motion detectors indicate the presence of motion adaptation.

Barlow and Hill (1963)\textsuperscript{35} were the first to show that responses of directional motion sensitive neurons decline gradually during prolonged motion stimulation. They suggested that this decline in response and subsequent reduction in background activity that occurred after the stimulus stopped, may be correlated with psychophysically measured motion after effects. Maddess and Laughlin (1985)\textsuperscript{33} reported that adaptation does not depend on velocity but rather on temporal frequency of the stimulus, and may relieve saturation.

de Ruyter van Steveninck (1986)\textsuperscript{36} found from his studies that value of the time constant depends on the magnitude of local velocity from which he concludes that the fly visual system uses estimates of local stimulus velocity to tune its filtering operations. Borst and Egelhaaf (1987)\textsuperscript{37} disagree with him over the origin of the adapting signal, suggesting that the adaptational state of a motion detector is governed mainly by the temporal frequency of the signal in its input channels.

Clifford and Langley (1995)\textsuperscript{38} in their adaptive Reichardt model, proposed that the adapting signal originates from locally integrated responses of a one dimensional array of elementary motion detectors, and is feedback to adapt the time constants of their temporal filters suggesting that adaptation to higher velocities is achieved by shortening the delay filter in the correlator model. Ibbotson (1998)\textsuperscript{39} tested this adaptive Reichardt model by recording from neurons in the wallaby visual system, found little change in the position of temporal frequency optimum before and after motion adaptation, despite showing changes to image step responses similar to those in fly neurons. Furthermore recent studies in the fly by Harris \textit{et al} (1999)\textsuperscript{11} shows little change in the temporal and spatial tuning properties of fly motion sensitive cells following adaptation, indicating that motion adaptation does not significantly alter the inherent velocity optimum of the EMDs. Instead this work reveals that motion adaptation induces a profound decrease in contrast sensitivity of fly motion sensitive cells via two proposed
mechanisms, a local direction sensitive after potential and a local direction insensitive contrast control. We propose that this gain reduction may serve to reduce sensitivity to image contrast, as well as reducing tendency of the motion detectors to saturate.

7. CONTRAST ADAPTATION

In the correlation model of motion detection, each correlator contains an expansive non-linearity (multiplication). This would make the correlator output particularly sensitive to the magnitude of input signals and so potentially vulnerable to saturation. The after potential and the gain reduction serve to release the motion pathway from this saturation, allowing to maintain a wide sensitivity across a wide range of stimulus conditions. The after potential acts antagonistically to recent activity in the cell, repositioning the cell’s responses within the available signalling range. By analogy with the retina, this type of subtractive mechanism may exploit correlations in continuous signals, reduce redundancy and maintain the operations of synapses in favorable regions of their input output functions.

Similarly the gain reduction component of adaptation scales down the magnitude of signals in the motion pathway. If the correlator non-linearity is to be protected from saturation the gain control should act on the inputs. Furthermore, since the output of the correlators depends on the spatio-temporal correlation between the input signals as well as their magnitude, the gain control would be best regulated by the magnitude of the correlator output, not the magnitude of the inputs. This is consistent with the observation that adaptation is recruited by motion but not flicker. The gain reduction mechanism will be recruited strongly when either the mean or the variance of the stimulus velocity distribution becomes large.

7.1. Experimental Verification

In order to test the relationships between image power spectra and velocity response curves, a set of experiments were carried out in which the recordings of steady-state responses of wide-field neurons in a hoverfly to motion of broad-band images at different velocities was taken and compared with the analytical and computational predictions. It was found that the shapes of these curves and their dependence on image statistics agree with theoretical predictions. Male specimens of the hoverfly *Volucella* are used for these experiments. The wide field neurons have proven particularly amenable to physiological analysis because their structure and physiological characteristics are nearly identical in different animals.

The results of the experiments suggested that while the system behaves like a simple Reichardt correlator at low contrasts, these curves support the presence of some form of contrast gain control at higher contrasts. This invariance with contrast is characteristic of the motion-adapted system; the unadapted system exhibits larger variations of the response level with stimulus contrast. From a practical point of view, the invariance of the motion-adapted velocity response curve with overall image contrast implies that mean correlator output may indeed provide an accurate estimate of velocity for a wide range of natural images.

8. CONTRAST GAIN REDUCTION - FEEDBACK ADAPTIVE EMD MODEL

In our elaborated Reichardt correlator array model, in order to remove the dependency of the response to changes in contrast and spatial frequency and to get an accurate estimate of velocity, contrast gain reduction is implemented by a feedback adaptive process as in Fig. 10. The gain of the EMD inputs is reduced by a signal derived from the rectified and low pass filtered outputs of a local EMD pool with different local preferred direction, fed back to control the gain of the EMD inputs. The model was inspired by recent observation of contrast dependent gain reduction in the responses of HS neurons following motion stimulation. The model captures several aspects of the adaptive phenomena observed in the biological system. In particular, the adaptation is strongest when the local motion detector output is the largest, conferring a robustness in adaptation to motion signals as opposed to static flicker or noise applied to the inputs. This matches data obtained from the electrophysiological experiments, which show that motion stimuli are much more effective at recruiting adaptive gain than other stimuli. Secondly, the adaptive mechanisms remain independent of the direction of local motion, despite the selectivity for motion as the source of adaptation.
Fig. 10. Block diagram of a feedback adaptive EMD model

Fig. 11. The input and the output responses of one EMD from the adaptive EMD array in response to the test sequence illustrated in Fig. 9(a).

In order to test the performance of this adaptive EMD array, we have compared the responses of the non-adapted elaborated EMD array with the adaptive elaborated EMD array at three different contrasts. From the graphs (Fig. 12 and Fig. 13) it is very clear that the adaptive feedback mechanism has resulted in decreasing the contrast sensitivity of the model, compared with an otherwise similar model that lacks adaptive feedback. It is also found that the adaptation is stronger at higher velocities and at higher contrasts.
Fig. 12. The correlator response of an unadaptive EMD array at three different contrasts. The response of the elaborated EMD array without including the adaptive feedback loop is noted at three different contrast (1, 0.5 and 0.1). It is clearly seen from the graph that there is huge variation of the response with contrast.

Fig. 13. The correlator response of an adaptive EMD array at three different contrasts. The response of the elaborated EMD array including the adaptive feedback loop is noted at three different contrast (1, 0.5 and 0.1). The adaptive feedback mechanism helps in decreasing the dependence of the response to contrast.

9. CONCLUSION AND FUTURE WORK

In this paper, we present a preliminary implementation of an adaptive EMD array model to take account of recent neurobiological research on the fly visual system. This model leads to contrast gain reduction at local
EMDs as observed in fly HS neurons. In our model, this gain reduction is achieved by direct feedback of local EMD outputs. An alternative mechanism in which the feedback signal gates a feedforward contrast normalisation provides an alternative explanation for the physiological data. Ongoing physiological experiments in our laboratory aim to test predictions of both models and thus to establish whether either of these models are fully consistent with motion adaptation mechanisms in the fly visual systems.

The present model demonstrates that a feedback gain control is capable of reducing the dependence of EMD output on contrast in natural images. Although the model does not in its present form, achieve ‘velocity constancy’ with respect to contrast, as implied by pilot experiments on the fly HS neurons, we have deliberately simplified this model to exclude known additional non-linearities of the fly vision, including logarithmic encoding of luminance and response saturation. Since earlier work already showed that these additional elaborations improve velocity coding by EMDs, inclusion of such components in future models may serve to achieve substantial improvements in performance of an adaptive EMD over earlier versions of the Reichardt correlator.

ACKNOWLEDGMENTS
This research is supported by funding from the US Air Force Office for Scientific Research / Asian Office for Aerospace Research & Development (contract # F62562-01-P-0158), the Sir Ross & Sir Keith Smith Fund and the Australian Research Council.

REFERENCES


