Steven D. Wiederman and David C. O'Carroll
Biologically inspired feature detection using cascaded correlations of off and on channels

© 2015. This work is licensed under the Creative Commons Attribution-NonCommercial-NoDerivatives 3.0 License. (CC BY-NC-ND 3.0)

Originally published at:
http://doi.org/10.2478/jaiscr-2014-0001

http://creativecommons.org/licenses/by-nc-nd/3.0/
BIOLOGICALLY INSPIRED FEATURE DETECTION USING CASCADED CORRELATIONS OF OFF AND ON CHANNELS

Steven D. Wiederman¹ and David C. O’Carroll¹

¹Adelaide Centre for Neuroscience Research, The University of Adelaide, South Australia, Australia, 5005
steven.wiederman@adelaide.edu.au, david.ocarroll@adelaide.edu.au

Abstract

Flying insects are valuable animal models for elucidating computational processes underlying visual motion detection. For example, optical flow analysis by wide-field motion processing neurons in the insect visual system has been investigated from both behavioral and physiological perspectives [1]. This has resulted in useful computational models with diverse applications [2,3]. In addition, some insects must also extract the movement of their prey or conspecifics from their environment. Such insects have the ability to detect and interact with small moving targets, even amidst a swarm of others [4,5]. We use electrophysiological techniques to record from small target motion detector (STMD) neurons in the insect brain that are likely to subserve these behaviors. Inspired by such recordings, we previously proposed an ‘elementary’ small target motion detector (ESTMD) model that accounts for the spatial and temporal tuning of such neurons and even their ability to discriminate targets against cluttered surrounds [6-8]. However, other properties such as direction selectivity [9] and response facilitation for objects moving on extended trajectories [10] are not accounted for by this model. We therefore propose here two model variants that cascade an ESTMD model with a traditional motion detection model algorithm, the Hassenstein-Reichardt ‘elementary motion detector’ (EMD) [11]. We show that these elaborations maintain the principal attributes of ESTMDs (i.e. spatiotemporal tuning and background clutter rejection) while also capturing the direction selectivity observed in some STMD neurons. By encapsulating the properties of biological STMD neurons we aim to develop computational models that can simulate the remarkable capabilities of insects in target discrimination and pursuit for applications in robotics and artificial vision systems.

1 Introduction

The dominant computational model for biological motion processing for over 50 years, the Hassenstein-Reichardt (HR) model for a direction-selective ‘elementary motion detector’ (EMD), involves correlation of spatially separated contrast signals after delaying one channel (Fig. 1A). The HR model (and a number of mathematically similar variants) is well supported by diverse evidence, from behavior to electrophysiology, particularly from insects [review, 1]. In dipteran flies, large arrays of local HR detectors with differing local direction preferences for motion are collated (i.e. spatially summed) by large-field neurons of the insect 3rd optic ganglion to provide matched filters for different classes of wide-field optical flow (e.g. pitch, roll, yaw and translation) and to mediate flight control [12].

While evidence for an HR-like detector in the wide-field motion system is overwhelming, less is known about the mechanisms underlying other types of motion processing in insects. This is partic-
ularly the case for the neural pathways involved in feature detection. One group of insect feature detectors, the small target motion detector (STMD) neurons, have several response characteristics that are well explained by classical HR models. These include distinct velocity tuning and a non-linear sensitivity to stimulus contrast [13]. Both direction-selective and non-directional STMDs with otherwise identical tuning have been described [9,14-16], yet the HR model is inherently selective for the direction of the stimulus. Moreover, the HR detector itself confers no selectivity for small targets—the hallmark of STMDs as a class and would thus be responsive to the larger features that dominate natural scenes [17]. Furthermore, the classical HR detector is insensitive to the sign of feature contrast, responding equally well to features that are brighter or darker than the background, while STMD neurons can be highly selective for a single sign of contrast (e.g. dark contrast targets).

To reconcile these disparate findings, we previously proposed a model for target selectivity by STMDs that utilized half-rectified separation of input channels for luminance increments (ON) and decrements (OFF) channels [6, 18]. We derived inspiration for this ‘elementary small target motion detector’ (ESTMD) model from electrophysiological recordings of rectifying, transient cells in the locust medulla [19], blowfly 1st optic chiasm [20] and blowfly medulla [6] that depolarize and/or spike (i.e. generate action potentials) in response to either luminance increments or decrements. Our ESTMD model included partially rectified ON and OFF channels with fast and independent dynamic adaptation (gain reduction), which are then correlated with one another after a delay at the same spatial location (Fig. 1B). This opposite polarity correlation exploits an expected property of a spatially circumscribed feature moving in a given direction: even against cluttered backgrounds, such a feature is most likely to have a leading edge opposite in sign to its trailing edge. Even a tiny object crossing the receptive field of a single detector will thus produce a response that first rises before falling. The addition of strong surround antagonism then extends this discrimination for small features along the axis orthogonal to its motion (Fig. 1B).

**Figure 1.** Correlation based models of ‘elementary motion detection’ (EMD). A. The Hassenstein-Reichardt (HR) EMD correlates spatially separated luminance signals with one path delayed in time before multiplication. The subtraction of mirror symmetric units permits opponent responses to motion in the non-preferred direction. B. The non-directional ESTMD model has strong centre-surround antagonism of ON and OFF channels followed by the correlation of the delayed OFF with the undelayed ON signal. This processing matches the spatiotemporal profile of a small moving target. C. In the 2 detector HR EMD, ‘like’ channels are correlated with one another, i.e. ON with ON and OFF with OFF. D. The cascaded EMD-ESTMD model uses spatiotemporally correlated ON and OFF channels to serve as inputs to an ESTMD model. F. The cascaded ESTMD-EMD model adds directionality to target detection by correlating the output of two non-directional ESTMD units.

Our ESTMD model accounts well for the size and velocity tuning observed in insect STMDs [6, 21], and is also robust to the presence of background clutter, even without relative motion cues [22]. This model therefore captures a remark-
able property of biological STMD neurons, which have been shown to respond very robustly to targets against highly cluttered backgrounds, regardless of the extent of relative motion [14, 15]. However the core correlation operation of the ESTMD model lacks the spatial asymmetry that confers direction selectivity to the HR-EMD model.

An interesting possibility to explain the observation that some STMD neurons are direction-selective is that they might employ a second-order motion detector that cascades an HR-type EMD stage with an ESTMD stage. This is further suggested by the recent discovery of direction-selective STMD neurons with relatively small receptive fields, around 5-10° across (where the inter-detector angle of the insect is 1-1.5°) [16]. While signaling local motion of small features, these small-field STMDs (SF-STMDs) still encapsulate inputs corresponding to a local pool of dozens of adjacent input ommatidia, so could easily represent the output of an asymmetric, non-linear integration of local ESTMDs – i.e. a 2nd order motion detector network [23].

Consistent with this notion, a 2nd order operation was recently proposed following our finding that the response time-course in the large-field dragonfly STMD neuron, CSTMD1, to small target stimuli builds to a maximum over several hundred milliseconds [10]. It does so, however, only for targets that successively track across tens of degrees of visual angle, and thus large numbers of local ommatidia [24, 25]. A second-order motion detector network [23] might not only confer direction selectivity, but could potentially enhance target detection by taking advantage of a distinguishing characteristic feature of natural target motion: true targets tend to move along continuous paths, even if they change direction or vary in contrast as they move across the background. A response in one local motion detector should be well correlated with an appropriately delayed response in neighboring detectors (i.e. matching the target velocity). Noise, on the other hand (including spurious feature motion of the background, such as foliage moving with wind), would tend to be local and inconsistent, and thus less likely to persist along continuous trajectories. A second-order system would thus enhance rejection of feature motion not correlated across multiple local adjacent input detectors, permitting amplification to enhance robustness whilst maintaining underlying selectivity to stimuli on the spatial scale of single ommatidia of the eye.

In this paper, we present several variants of computational models that capture the key properties of direction-selective SF-STMD neurons by cascading partially rectified ESTMD stages with HR-type EMDs, and vice versa. We further extend the work presented in Wiederman & O’Carroll, 2013b [26], with the addition of new physiological data from STMD neurons (direction selectivity, contrast sensitivity and dark target selectivity). Additionally, we extend model simulations, robustly determining the contrast sensitivity function of a particular STMD neuron (CSTMD1). We examine the key response tuning and predictions of these second order systems for simple stimuli. We find that the combination of an ESTMD front-end with a 2nd order EMD is able to combine both the direction selectivity of classical EMDs with the feature selectivity of the ESTMD model. This model captures the sharper response tuning for either size or velocity of targets observed in insect STMD neurons compared with the predictions of either EMD or ESTMD models alone.

2 Results
2.1 Model variants

We developed three versions of the target-detection model each including identical ‘biomimetic’ front-end processing. The computational models simulated insect optics (blurring and hexagonal sampling) and early visual processing (dynamic band pass filtering) as described in the Appendix [6-8]:

i ESTMD: this is our original ESTMD model (Fig. 1B) which includes strong surround antagonism and correlates a delayed OFF channel with an undelayed ON channel [6, 18].

ii 2nd order EMD-ESTMD: this is a second order ‘hybrid’ model (Fig. 1D). This implements a front end based on a direction-selective 2-detector EMD (Fig 1C) that separates ON and OFF channels and then subsequently correlates these ‘like’ channels in separate multiplication stages. This EMD is based on recent findings
from genetic modification of fruit-fly vision and is currently regarded as the ‘state of the art’ as a biomimetic model for the HR-EMD as observed in real insects [28]. In our hybrid EMD-ESTMD model (Fig. 1D) each ‘ON-ON’ and OFF-OFF motion signal at the output of the 2 EMDs then serves as input to a 2nd order ESTMD stage which correlates the delayed signal from the OFF EMD with the undelayed ON EMD.

iii 2nd order ESTMD-EMD: this variant (Fig. 1E) implements similar operations to the EMD-ESTMD, but in reverse order, i.e. the model first computes the ‘matched target’ filter with an ESTMD (as in Fig. 1B), before correlating these as inputs to a 2nd order 2-detector EMD (as in Fig. 1C).

2.2 Direction selectivity

A subset of STMD neurons are direction selective [16]. As shown in Fig. 2A, an example STMD neuron responds robustly to targets drifted across the display from right to left (at varying vertical locations), but not to targets drifted in the rightwards direction. In this example, responses are in fact inhibited below the spontaneous firing rate to such an ‘opponent’ direction of motion. In a second example STMD, responses to targets drifted in either direction produce robust spiking activity (i.e. non-directional). Such ‘receptive fields’ of the neuron can also be mapped in the vertical direction, to determine the neuron’s overall preferred directionality (data not shown).

Fig. 3 shows that both hybrid models (Fig 1D, E) lead to a direction-selective output. Shown is the percentage strength of directionality, defined here relative to the sum of the absolute responses to targets moving in two opposite directions. The target stimulus subtended an angle of $1.25^\circ \times 2^\circ$ and moved at a speed of $45^\circ/s$. As can be seen, the ESTMD responds equally to either target direction and is in effect a local flicker detector, matched to the spatiotemporal profile of a moving target. A cascaded 2-detector EMD either before (EMD-ESTMD) or after (ESTMD-EMD) the matched filter induces strong directionality to the model responses, with responses in the ‘preferred’ direction comprising almost 100% of the overall response to both.

2.3 Spatiotemporal tuning

To examine whether model variants produce similar characteristics to those observed in physiological STMD experiments [9, 27], we examined the target size tuning of the model variants by simulating target motion of varying width (i.e. their spatial direction in the direction of travel) and at two different velocities (Fig. 4).

All 3 model variants produce similar response characteristics, displaying a dependence on both image velocity and target width. These tuning functions are similar to those observed in earlier ESTMD modeling efforts [28].

![Figure 2](image2.jpg)

**Figure 2.** The receptive field of a neuron is mapped by drifting a target across the screen at multiple vertical locations and plotting the spiking activity A. An unidentified STMD neuron is only responsive to a target drifting in one direction (left). B. The receptive field of CSTMD1 reveals that the neuron responds similarly to target motion either left or right (non-directional).
2.2 Direction selectivity

A cascaded 2-detector EMD either before (EMD-ESTMD) or after (ESTMD-EMD) the matched filter induces strong directionality in the response to a moving target. ESTMD responds equally to either target direction and is in effect a local flicker detector, matched responses in the ‘preferred’ direction (ESTMD-EMD) or after (ESTMD-EMD) the matched filter. As can be seen, the neuron responds robustly to targets drifted across the display from right to left (at varying vertical locations).

The tuning curve to drifting targets at velocities 45°/s exhibits a similar optimum to the ESTMD model, however is also more sharply tuned. At a lower, less optimal velocity (45°/s), curves are slightly left-shifted compared to a velocity of 90°/s, indicating the confounded relationship between the velocity/width profile of a moving target.

2.4 Responses to ON and OFF edges

An interesting prediction of the ESTMD model (irrespective of the addition of a second order correlation) that has not previously been extensively tested in the biological STMDs is that responses to individual ON or OFF edges should be smaller than responses to a dark target (due to the multiplicative nonlinearity between the opposite polarity ON and OFF contrast channels).

The tuning curves of the three model variants at two velocities: 45°/s (closed circles) and 90°/s (open circles) shows tuning curves similar to those observed in physiology. The EMD-ESTMD model produces a tuning curve shifted to the right in comparison to the other models. The ESTMD-EMD model has a similar optimum to the ESTMD, however is more sharply tuned.

The EMD-ESTMD size-tuning curve is sharper and shifted to the right (i.e. less selective for very small targets) than with the other models. The ESTMD-EMD tuning curve exhibits a similar optimum to the ESTMD model, however is also more sharply tuned. At a lower, less optimal velocity (45°/s), curves are slightly left-shifted compared to a velocity of 90°/s, indicating the confounded relationship between the velocity/width profile of a moving target.

The EMD-ESTMD size-tuning curve is sharper and shifted to the right (i.e. less selective for very small targets) than with the other models. The ESTMD-EMD tuning curve exhibits a similar optimum to the ESTMD model, however is also more sharply tuned. At a lower, less optimal velocity (45°/s), curves are slightly left-shifted compared to a velocity of 90°/s, indicating the confounded relationship between the velocity/width profile of a moving target.

Fig. 5 shows data obtained from a dragonfly STMD neuron that suggest that this untested characteristic is indeed observed in the biological system. Interestingly, it appears in this example that the dark target responses represent a super-linear combination of the responses to ON and OFF ‘edge’ features produced by extending a single bar across the stimulus display, along the same path as the target (as indicated by the inset pictograms). This observation is strongly supportive of an ESTMD-like operation occurring in the input pathway to the STMD neuron.
Figure 6. Neuron and Model output in response to moving stimuli of varying contrast. A. Compares the contrast sensitivity function (CSF) for responses to sinusoidal grating patterns of optimum frequency in HS cells, a class of neurons in the fly that are believed to collate the output from large arrays of classical HR-EMDs [29], with those of CSTMD1 (a dragonfly STMD neuron selective for small targets) to optimal target stimuli. Because the wide-field neuron integrates motion from many local EMDs, its absolute sensitive to low contrast motion is higher than in the STMD, hence our use of a Log X-axis to display the data. However once above a threshold level, the STMD displays a much steeper rise of response towards the saturated regime. B. Shows the sensitivity of the 3 model variants to Weber contrast \( \left( \frac{I_{\text{target}}-I_{\text{background}}}{I_{\text{background}}} \right) \), a measure that is more appropriate to small targets against a uniform background. Although all 3 variants show an expansive response to contrast, the EMD-ESTMD and ESTMD-EMD models exhibit similar CSFs with a higher-order dependence on contrast (i.e. steeper rise) than the ESTMD, due to the 4th order (cascaded) multiplicative correlations. C. We apply an arbitrary gain to each of the model variants, matching the initial expansive ‘threshold’. We also include a saturating non-linearity (a tanh function), ubiquitous in neuron encoding. Note that the expansiveness from either model variant (single correlation or cascade) is readily fit to the physiological data by varying the arbitrary gain and saturation parameters.

2.5 Responses to feature contrast

A further characteristic of the HR EMD is that output in response to increased stimulus contrast is expansive, due to the super-linear interaction between the inputs (usually modeled as multiplication). In biological systems, this expansive non-linearity is inherently bounded by the saturation limits of synaptic signaling and biochemistry within which such super-linear operations must be implemented [1].

Hence typical neurons taking their inputs from HR-type EMDs, such as the HS neurons of flies, show contrast-response functions that are initially super-linear [30], before becoming sub-linear as saturation begins to dominate the response (Fig.
6A). If we model the ESTMD with no such saturation on its outputs (Fig. 6B), it displays an identical expansive characteristic, due to the multiplication stage. The EMD-ESTMD and ESTMD-EMD show very similar responses to increasing target contrast. However, as expected from cascading a 2\textsuperscript{nd} order multiplication, the rate of expansion of both is even higher order (i.e. 4\textsuperscript{th} order) than in the ESTMD.

Can the shape of the contrast response function reveal whether such cascaded multiplications underlie physiological responses? Unfortunately, this is unlikely, as arbitrary sensitivity (gain) with the addition of a saturating nonlinearity (a fundamental component of a neurons response characteristics) could be used to fit either model variant (Fig. 6C).

3 Discussion

Several alternative correlation-based models (including the ESTMD) can explain a number of basic response tuning properties of insect STMD neurons, such as their size and velocity tuning. Other characteristics, such as selectivity for contrast polarity (e.g. dark) and their ability to discriminate targets in clutter without relative motion cues, can presently only be explained by the ESTMD model [6-8]. However, some STMDs also exhibit characteristics that the ESTMD cannot account for, including directionality [9] and facilitation [10]. By elaborating the ESTMD model with a cascaded EMD, we induce directionality (Fig. 2, 3) maintaining the size and velocity tuning intrinsic to our ESTMD model (Fig. 4) and can also explain the highly super-linear response to target contrast that distinguishes many insect STMD neurons (Fig. 5, 6).

A 2\textsuperscript{nd} order motion detector network similar to those we implement was previously proposed to explain the ‘facilitated’ responses observed in some STMDs when targets move along continuous trajectories [10, 24, 25]. In future experiments we will present moving targets in cluttered environments to our model variants and determine whether this facilitation breaks or enhances target discrimination.

**Figure 7.** Model Overview of A. early visual processing and B. the elementary small target motion detector (ESTMD). Inputs are spatially blurred to represent fly optics. We retain only the green channel of the RGB image, to represent spectral sensitivity of the motion pathway. A complex photoreceptor model implements dynamic filtering characteristics and adaptive feedbacks which allow for the encoding of vast luminance conditions. LMCs are modeled as dynamic spatiotemporal high-pass filters (relaxed), removing redundant information. The model implements functionality inspired from electrophysiological recordings of RTCs, found in the brain of the fly. This includes ON and OFF channel separation, independent fast temporal adaptation and independent channel surround antagonism. Finally, the delayed OFF channel is recombined with the undelayed ON channel for dark target sensitivity. The final output reveals enhanced small target discrimination as seen in physiological STMDs.

One of our objectives was to determine whether one or other of the cascaded model variants matched observed physiological results and whether in fact one variant could be ‘ruled out’ due to an inconsistency. As both models match the data we have supported the existence of a 2\textsuperscript{nd} order network, however, more physiological experiments and modeling will be required to elucidate the underlying architecture. Interestingly, it has recently been shown that insects respond behaviorally to non-Fourier motion [31], a further indication of a 2\textsuperscript{nd} order network underlying feature discrimination. In future work we will stimulate our model variants with non-Fourier motion and see whether...
this can be used to determine which architecture is the more likely candidate to underlie STMD processing.

4 Appendix

4.1 Model Overview

Models were implemented in Matlab (Mathworks, Natick, USA). An overview of early visual processing and the ESTMD model is shown in Fig. 7. The EMD model and combination model variants are shown in Fig. 1.

Photoreceptor responses were based on a biomimetic model with parameters and elaborations derived via electrophysiological results from *Eristalis tenax* [32]. The spatiotemporal dynamics of the first order interneurons, the LMCs, have previously been established for blowfly (*Calliphora vicina*) and hoverfly, (*Eristalis tenax*) [33, 34]. Model outputs were matched to STMD neurons in *Eristalis tenax* [14-16]. We modeled an array of ESTMD (elementary STMD) subunits that could be spatially pooled to form the position invariant receptive field, as seen in the physiological STMD [27]. The model output define values in a three dimensional (2D+t) space.

4.2 Early Visual Processing

We used the green channel in our 8-bit input imagery to simulate 'green-blue' spectral sensitivity in the fly visual system [35]. We applied a Gaussian low-pass filter (full width at half maximum 1.4°) to emulate spatial blur of the optics [36]. Input images were spatially sampled at 1° in a hexagonal manner [37]. Photoreceptors incorporated variable gain control, saturating nonlinearities and dynamic low-pass filtering [32, 38], with cutoff frequencies dependent on adaptation state (ranging from 20 to 100 Hz). Followed by two divisive, delayed feedbacks (one linear, one exponential) representing short and longer term adaptations (τ=23 ms, τ=12.4 s). Finally, a compressive, saturating nonlinearity is implemented by a Naka-Rushton transform. The LMC implements spatiotemporal high-pass filtering, altering its filtering characteristics dependent on visual conditions. In the dark adaptation state, the LMC is more integrative with longer sustaining temporal components.

As overall luminance conditions increase, the LMC becomes more transient and high-pass in nature, both in space and time [33]. This spatial interaction (center-surround antagonism) was modeled in a feed forward manner with surround (nearest neighbor) photoreceptor signals summed, and temporally delayed (τ=16 ms), before subtractive inhibiting the central LMC (strength ranging from 0-30%, dependent on adaptation state). The LMC temporal dynamics were modeled with relaxed, variable high-pass filtering (lower corner frequency ranging from 0 to 8 Hz) which incorporated a small DC component (0-10%). Following, a saturating nonlinearity with a hyperbolic tangent function, ensured the LMC response was limited to a predictable output range.

4.3 Elementary Small Target Motion Detector (ESTMD)

The main processing of the ESTMD is based on the Rectifying Transient Cell (RTC). Briefly, the RTC creates transient ‘on’ and ‘off’ phases (from the LMC high-pass filtering), separated via a further temporal high pass filter (τ = 40ms) and then half-wave rectification into independent ON and OFF channels. Each of the channels is temporally processed through a fast adaptive mechanism. An adaptation state is determined by a nonlinear filter, which approximates cellular ‘fast depolarization and slow repolarization’ responses. This low-pass filter switches its time constant dependent on whether the input is increasing or decreasing (time constants are ‘fast’ (τ=3 ms) when channel input is increasing and ‘slow’ (τ=70 ms) when decreasing). This adaptation state causes subtractive inhibition of the unaltered ‘pass-through’ signal. The result of this complex, nonlinear filtering is the signaling of ‘novel’ transient contrast changes (of the particular channel phase, ‘on’ or ‘off’) with the suppression of fluctuating textural variations. As well as this temporal antagonism, the channels also exhibit spatial antagonism with ON surround channels subtractive inhibition of the ON center channel, and similarly with the OFF channels. The resultant signal was then half-wave rectified, so that the surround does not inhibit the center below a zero value (a nonlinearity seen in some spiking neurons).

ON and OFF channels are recombined via multiplication [6]. Both dark and light target sensitivity
is possible by delaying and recombining the relevant contrast polarity. For these experiments we delayed the OFF channel using a 1st-order low-pass filter ($\tau = 25$ ms) and multiplying this by the undelayed ON channel. This processing provides a template for the characteristic temporal 'signature' of a small moving target.

### 4.4 2 Detector Elementary Motion Detector (EMD)

Channels were separated into ON and OFF channels via half-wave rectification following high pass filtering ($\tau = 100$ ms). The delay arm was modeled with a low-pass filter ($\tau = 25$ ms).

## 5 Acknowledgment

Research supported by the Australian Research Council’s Discovery Projects funding scheme (project number DP130104572) and the US Air Force Office of Scientific Research (FA2386-10-1-4114). We thank the manager of the Botanic Gardens in Adelaide for allowing insect collection.

## References


