

A Neurobiological and Computational
Analysis of Target Discrimination in Visual
Clutter by the Insect Visual System

Steven Wiederman

Discipline of Physiology

School of Molecular & Biomedical Science

The University of Adelaide

September 2008

A thesis submitted for the degree of Doctor of Philosophy

THESIS ABSTRACT

Some insects have the capability to detect and track small moving objects, often against cluttered moving backgrounds. Determining how this task is performed is an intriguing challenge, both from a physiological and computational perspective. Previous research has characterized higher-order neurons within the fly brain known as ‘small target motion detectors’ (STMD) that respond selectively to targets, even within complex moving surrounds. Interestingly, these cells still respond robustly when the velocity of the target is matched to the velocity of the background (i.e. with no relative motion cues).

We performed intracellular recordings from intermediate-order neurons in the fly visual system (the medulla). These full-wave rectifying, transient cells (RTC) reveal independent adaptation to luminance changes of opposite signs (suggesting separate ‘on’ and ‘off’ channels) and fast adaptive temporal mechanisms (as seen in some previously described cell types). We show, via electrophysiological experiments, that the RTC is temporally responsive to rapidly changing stimuli and well suited along a proposed pathway to target-detecting neurons.

To model this target discrimination, we use high dynamic range (HDR) natural images to represent ‘real-world’ luminance values that serve as inputs to a biomimetic representation of photoreceptor processing. Adaptive spatiotemporal high-pass filtering (1st-order interneurons) shape the transient ‘edge-like’ responses, useful for feature discrimination. Following this, a model for the RTC implements a nonlinear facilitation between the rapidly adapting, and independent polarity contrast channels, each with center-surround antagonism. The recombination of the channels results in increased discrimination of small targets, of approximately the size of a single pixel, without the need for relative motion cues. This method of feature discrimination contrasts with traditional target and background motion-field computations. We show that our RTC-based target detection model is well matched to properties described for the higher-order STMD neurons, such as contrast sensitivity, height tuning and velocity tuning. The model output shows that the spatiotemporal profile of small targets is sufficiently rare within natural scene imagery to allow our highly nonlinear ‘matched filter’ to successfully detect many targets from the background.

The model produces robust target discrimination across a biologically plausible range of target sizes and a range of velocities. We show that the model for small target motion detection is highly correlated to the velocity of the stimulus but not other background statistics, such as local brightness or local contrast, which normally influence target detection tasks.

From an engineering perspective, we examine model elaborations for improved target discrimination via inhibitory interactions from correlation-type motion detectors, using a form of antagonism between our feature correlator and the more typical motion correlator. We also observe that a changing optimal threshold is highly correlated to the value of observer ego-motion. We present an elaborated target detection model that allows for implementation of a static optimal threshold, by scaling the target discrimination mechanism with a model-derived velocity estimation of ego-motion.

Finally, we investigate the physiological relevance of this target discrimination model. We show that via very subtle image manipulation of the visual stimulus, our model accurately predicts dramatic changes in observed electrophysiological responses from STMD neurons.

AUTHOR'S STATEMENT

This work contains no material which has been accepted for the award of any other degree or diploma in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text. I give consent to this copy of my thesis when deposited in the University Library, being made available for loan and photocopying, subject to the provisions of the Copyright Act 1968.

The author acknowledges that copyright of published works contained within this thesis (as listed below) resides with the copyright holder(s) of those works.

(1) Wiederman S.D., Shoemaker P.A., O'Carroll D.C. (2008) A model for the detection of moving targets in visual clutter inspired by insect physiology. PLoS ONE 3(7): e2784 doi:10.1371/journal.pone.0002784

Copyright: © 2008 Wiederman et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

(2) Wiederman, S., Shoemaker, P. and O'Carroll, D.C. (2007) Biologically inspired small target detection mechanisms. Proceedings IEEE, Intelligent Sensors, Sensor Networks and Information Processing (ISSNIP) Conference, pp 269 - 275

Copyright: © 2007 Institute of Electrical and Electronics Engineers, Incorporated (the "IEEE") all rights reserved

(3) Wiederman, S.D., Brinkworth, R.S.A. and O'Carroll, D.C. (2008) Bio-inspired small target discrimination in high dynamic range natural scenes. IEEE Proceedings, Bio-inspired Computing: Theories and Applications (BIC-TA), in press.

Copyright: © 2008 Institute of Electrical and Electronics Engineers, Incorporated (the "IEEE") all rights reserved

(4) Wiederman, S.D., Brinkworth, R.S.A. and O'Carroll, D.C. (2008) Bio-inspired target detection in natural scenes: optimal thresholds and ego-motion, in Biosensing, Proceedings of the SPIE Vol.7035, 70350Z

Copyright: © 2008 Institute Society of Photo-Optical Instrumentation Engineers (SPIE) all rights reserved

SIGNATURE _____ DATE _____

ACKNOWLEDGEMENTS

Most importantly, I would like to thank Alicia, who moved with me from Sydney so that I could join the O'Carroll lab in Adelaide. Without her love and support, none of this would have been possible.

I would like to thank my supervisor Dave O'Carroll, who knows the perfect balance between being mentor, friend and family support. I have learned so much these last few years, all of which was due to his great guidance and helpful encouragement.

Thanks to Paul Barnett, for being a fantastic office buddy and good friend – it was a pleasure to share the PhD experience with you.

Thank to Russell Brinkworth and Karin Nordström, both for your friendship and help with my research. I was very lucky, as I was able to learn from each one of you, entirely different things, and it has made me an 'all round' better scientist and engineer.

Thanks to my external supervisors; to Subhash Challa for setting me up here in Adelaide, and to Pat Shoemaker for his help in my project and our work together.

Many thanks to the O'Carroll lab family, Doug Bolzon, Jodi Gray, Ximena Nelson, Bob Piddington, Bart Geurten, Irene Moyer and Eng Leng Mah. Also, thanks to the extended family, everyone in the Discipline of Physiology.

Statement of Contributions of Jointly Authored Works

The following states the contribution of the authors to the following published works.

Steven Wiederman	SDW	Patrick Shoemaker	PAS
Russell Brinkworth	RSAB	David O'Carroll	DO'C

(1) Wiederman S.D., Shoemaker P.A., O'Carroll D.C. (2008) A Model for the Detection of Moving Targets in Visual Clutter Inspired by Insect Physiology. PLoS ONE 3(7): e2784 doi:10.1371/journal.pone.0002784

Conceptualisation: SDW wished to understand the functionality of 'on-off' units, whether as non-directional motion detectors, inputs to correlator motion detectors, or feature detectors. Upon discussion with DO'C, SDW agreed to model these units as contrast gain controllers in the arms of motion correlators (poster presentation), whilst SDW discussed the relevance of 'on-off' units as feature detectors and as possible inputs to motion detectors with DO'C and PAS. SDW and PAS suggested that the presence of ON- and OFF-signaling pathways with independent adaptation might well be relevant for small feature extraction (by the novel on and off contrast phases (independent and unadapted) 'hitting' of the small subunits within the 'on-off' unit). PAS's immediate interest was in small target detection which was reinforced by the presence of inhibitory surrounds in the receptive fields of these cells (as indicated by the previous literature). SDW conceptualized elaborated models for this spatial filtering. These various ideas coalesced into a suitable nonlinear matched filter for small moving targets. This required a model to test for this target discrimination which was then designed and implemented in Simulink/Matlab by SDW.

Realisation: All electrophysiological investigation undertaken by SDW, under the tutelage of DO'C. Modeling: Lipetz function PAS, with varying midpoint parameter by SDW. Original model for fast depolarisation, slow repolarisation adaptive mechanism developed by PAS. Implementation of nonlinear filter was provided in SPICE by PAS that coincidentally was mathematical equivalent to gradient 'switched' filter designed by SDW. All simulation and analysis tools written by SDW. Suggested use of natural scenes and panoramic inputs by DO'C. ROC analysis suggested by PAS. Examining temporal responsiveness of RTC by DO'C, SDW and PAS with PAS suggesting 'doublet' stimulus. All Python coding of new visual stimuli by SDW. All electrophysiological data analysis and model analysis by SDW.

Documentation: First writing by SDW, with both PAS and DO'C providing strong written contributions and review.

(2) Wiederman, S., Shoemaker, P. and O'Carroll, D.C. (2007) Biologically Inspired Small Target Detection Mechanisms. Proceedings, Intelligent Sensors, Sensor Networks and Information Processing (ISSNIP) Conference, pp 269 – 275

Conceptualisation and Realisation: this is an engineering publication of the modelling results as seen in the PLoS research, therefore see (1) above.

Documentation: written by SDW

(3) Wiederman, S.D., Brinkworth, R.S.A. and O'Carroll, D.C. (2008) Bio-inspired small target discrimination in high dynamic range natural scenes. Proceedings, Bio-inspired Computing: Theories and Applications (BIC-TA), in press.

Conceptualisation: Idea to test model across varying target sizes and velocities by SDW. Interest in seeing capability for target discrimination with HDR stimulus (due to photoreceptor dynamics) by RSAB and SDW. Idea to test for correlations with image luminance and contrast values, and panorama velocities by SDW.

Realisation: Translation of overall modelling efforts to Matlab was a collaboration by SDW and RSAB. All elaborated photoreceptor modelling by RSAB. Extended modelling of LMC dynamics by SDW and RSAB. All analysis of results by SDW.

Documentation: written by SDW, with RSAB providing editing and review. Useful suggestions by D'OC.

(4) Wiederman, S.D., Brinkworth, R.S.A. and O'Carroll, D.C. (2008) Bio-inspired target detection in natural scenes: optimal thresholds and ego-motion, in Biosensing, Proceedings of the SPIE, Vol 7035, 70350Z

Conceptualisation: The model elaborations were conceptualised by SDW. Insightful mathematical interpretation and helpful analysis of ideas provided by RSAB.

Realisation: Target discrimination model design, implementation and analysis by SDW. All EMD/Wide-field components of the model written by RSAB. Interactions between the two written by SDW. All analysis of results by SDW.

Documentation: written by SDW, with RSAB providing editing and review Useful suggestions by D'OC.

I agree with the above statement of contribution and give permission for the paper(s) to be included in this thesis.

Steven Wiederman: SIGNATURE _____ DATE _____

Patrick Shoemaker: SIGNATURE _____ DATE _____

Russell Brinkworth: SIGNATURE _____ DATE _____

David O'Carroll: SIGNATURE _____ DATE _____

AUTHOR'S COMMENTS

- Publications within this thesis are exact copies of the original article except for the following three considerations
 1. Typesetting has been changed to provide a consistent format across the chapters of the thesis.
 2. References at the end of each article have been removed and can be found at the end of the thesis.
 3. Where appropriate, figure cross-referencing has been altered.
 4. A couple of titles have been added to figures within the publications, to provide a clearer Table of Figures.
- Within journal articles, figures are cross referenced as they were published, e.g. figure 4 in Chapter 2 is captioned 2.4, but referenced in text as Figure 4.
- There are variations in spelling between the chapters, between Australian and United States versions (due to submission rules for article publications).

TABLE OF CONTENTS

Chapter 1: Introduction	17
1.1 Target Detection, Tracking and Pursuit.....	17
1.2 Motion Detection Theory.....	19
1.2.1 Feature-tracking.....	19
1.2.2 The Gradient Scheme.....	19
1.2.3 The Reichardt Correlator	21
1.2.4 The Motion Stimulus (Classical and Natural Images).....	23
1.3 Insect Visual Pathway.....	25
1.3.1 Optics.....	25
1.3.2 Photoreceptors	26
1.3.3 Neural Superposition	26
1.3.4 Lamina Neurons.....	27
1.3.5 On-Off cells	30
1.3.6 The Neural Basis for an Elementary Motion Detector (EMD).....	31
1.4 Neural Mechanisms and Models for Motion Detection.....	33
1.4.1 Lobula Plate Tangential Cells (LPTC)	33
1.4.1.1 Figure Detection (FD) neurons	33
1.4.2 LGMD/DCMD.....	34
1.4.3 Target selective neurons in the insect brain.....	35
1.4.3.1 Small Target Motion Detectors (STMD)	35
1.5 Neuromorphic Models for Target Selectivity.....	37
1.5.1 Photoreceptors	37
1.5.2 Figure Detection (FD) Cells	37
1.5.3 Elaborated FD.....	38
1.5.4 STMD	39

Chapter 2: Rectifying Transient Cells and Target Discrimination.....	40
2.1 Context.....	40
2.2 A Model for the Detection of Moving Targets in Visual Clutter Inspired by Insect Physiology.....	42
2.2.1 Abstract.....	42
2.2.2 Introduction.....	43
2.2.2.1 Computational models for target discrimination.....	44
2.2.2.2 Full-wave rectifying transient neurons	44
2.2.2.3 Rectifying Transient Cells in the target detection pathway	47
2.2.3 Methods.....	47
2.2.3.1 Modeling.....	47
2.2.3.2 Input imagery	48
2.2.3.3 Physiology	51
2.2.4 Results.....	52
2.2.4.1 Photoreceptors.....	52
2.2.4.2 Large Monopolar Cells	53
2.2.4.3 Rectifying Transient Cell.....	54
2.2.4.4 Comparison of model responses to fly RTCs	56
2.2.4.5 Temporal responsiveness	57
2.2.4.6 Contrast Sensitivity Function.....	58
2.2.4.7 Target height tuning.....	60
2.2.4.8 Velocity tuning.....	61
2.2.4.9 Responses to targets in clutter.....	63
2.2.4.10 Relative motion.....	67
2.2.5 Discussion	68
2.2.6 Conclusion	69

2.2.7 Acknowledgements.....	69
Chapter 3. Engineered Target Discrimination Derived from Insect Physiology..	70
3.1 Context.....	70
3.2 Biologically Inspired Small Target Detection Mechanisms	71
3.2.1 Abstract.....	71
3.2.2 Introduction.....	72
3.2.3 Modelling & Simulation	74
3.2.3.1 Overview	74
3.2.3.2 Input Images.....	75
3.2.3.3 Early Visual Processing	75
3.2.3.4 Rectifying Transient Cell	76
3.2.4 Results.....	78
3.2.5 Discussion.....	79
3.2.6 Acknowledgement	81
Chapter 4: Target Discrimination in General Stimulus Conditions.....	82
4.1 Context.....	82
4.2 Bio-inspired Small Target Discrimination in High Dynamic Range Natural Scenes	84
4.2.1 Abstract.....	84
4.2.2 Introduction.....	85
4.2.3 Methods (Modeling & Simulation).....	86
4.2.3.1 Overview	86
4.2.3.2 Input Images.....	87
4.2.3.3 Photoreceptors.....	89
4.2.3.4 Large Monopolar Cells	90
4.2.3.5 Rectifying Transient Cells	90

4.2.3.6 Simulation and Analysis	93
4.2.4 Results & Discussion	94
4.2.5 Conclusion	100
4.2.6 Acknowledgements.....	101
Chapter 5: Elaborated Models for Target Detection.....	102
5.1 Context.....	102
5.2 Bio-inspired Target Detection in Natural Scenes: Optimal Thresholds and Ego-motion	103
5.2.1 Abstract.....	103
5.2.2 Introduction.....	104
5.2.3 Modeling, Simulation & Analysis	107
5.2.3.1 Model Overview	107
5.2.3.2 Bio-inspired & Biomimetic.....	107
5.2.3.3 Input Images.....	108
5.2.3.4 Early Visual Processing	109
5.2.3.5 Rectifying Transient Cells	110
5.2.3.6 Elementary Motion Detectors	111
5.2.3.7 Near-local EMD inhibition	113
5.2.3.8 Ego-motion scaling	115
5.2.4 Conclusion	118
5.2.5 Acknowledgements.....	120
Chapter 6: Testing the ESTMD Model with Experiments on STMDs	121
6.1 Context.....	121
6.2 Comparison of the ESTMD Model to the Dragonfly CSTMD Responses.....	122
6.2.1 Introduction.....	122
6.2.2 Methods.....	122

6.2.2.1 Electrophysiology	122
6.2.2.2 CSTMD Identification	123
6.2.2.3 Number of Trials	123
6.2.2.4 Image preparation	123
6.2.3 Results.....	124
6.2.3.1 Spiking Receptive Field Mapping.....	124
6.2.3.2 Graded Receptive Field Mapping	126
6.2.3.3 Contrast Polarity Selectivity	127
6.2.3.4 On and Off Channels in the Insect Visual System.....	129
6.2.3.5 Natural Image Experiments.....	130
6.2.3.6 A Note on Space and Time	133
6.2.3.7 CSTMD Electrophysiological Responses to the Natural Images.....	134
6.2.4 Discussion.....	136
Chapter 7: Conclusions	140
References.....	145

TABLE OF FIGURES

Figure 1-1. Fly chase behaviour.....	18
Figure 1-2. The Gradient Detector.....	20
Figure 1-3. The Reichardt Correlator.....	21
Figure 1-4. Sexual Dimorphism.....	25
Figure 1-5. Fly Eyes.....	27
Figure 1-6. The Lamina and Medulla of <i>Drosophila</i>	29
Figure 1-7. Cross-section of the Fly Head.....	31
Figure 1-8. Photoreceptor Processing.....	37
Figure 1-9. Small Field FD model.....	38
Figure 2-1. Model Overviews.....	46
Figure 2-2. Input Images and Optical Blurring.....	50
Figure 2-3. Panorama Rotation.....	53
Figure 2-4. Rectifying Transient Cells (Independent Adaptation).....	56
Figure 2-5.. Temporal Responsiveness.....	58
Figure 2-6. Contrast Sensitivity Function (Small Target and Wide-field).....	59
Figure 2-7. Target Height Tuning.....	61
Figure 2-8. Velocity Tuning.....	62
Figure 2-9. Sample Data Traces.....	64
Figure 2-10.. Natural Image Panoramas (ROC curves).....	66
Figure 2-11. Relative Motion.....	67
Figure 3-1. Responses of Model Rectifying Transient Cell.....	73
Figure 3-2. Overview of the ESTMD model.....	74

Figure 3-3. An example of model outputs.	77
Figure 3-4. ROC curves for four panoramic images.	79
Figure 4-1. Overview of the ESTMD model.	88
Figure 4-2. Complex Photoreceptor Model.	90
Figure 4-3. HDR Panoramic Images.....	92
Figure 4-4. Model output stages.	95
Figure 4-5. Receiver Operating Characteristic analysis.	96
Figure 4-6. Target Size AUROC.	98
Figure 4-7. Panorama Velocity AUROC.....	99
Figure 4-8. Model Response correlations.	100
Figure 5-1. Elaborated Model Overview.	108
Figure 5-2. Output of the elaborated ESTMD model.	112
Figure 5-3. ESTMD with inhibition.	114
Figure 5-4. Ego-motion rescaling technique.....	116
Figure 5-5. Ego-motion rescaled ESTMD.	117
Figure 5-6. Cumulative Distribution Functions.....	118
Figure 6-1. Single Scan through Receptive Field.....	124
Figure 6-2. Spiking Receptive Field Scan.	125
Figure 6-3. High Contrast Sensitivity.	126
Figure 6-4. Graded Receptive Field Scan.....	127
Figure 6-5. Dark Target Selectivity of CSTMD.....	128
Figure 6-6. Target Height Tuning (Dark and Light Polarity).....	129
Figure 6-7. An Off channel in the medulla of the blowfly, <i>Calliphora</i>	130
Figure 6-8. ESTMD Model Responses to the Natural Images.	132

Figure 6-9. Mirrored Receptive Field. 133

Figure 6-10. CSTMD Neural Responses to the ‘Bush’ Images. 135

Figure 6-11. CSTMD Responses to the ‘Car Park’ Images. 136

Figure 6-12. Morphology of CSTMD. 138

CHAPTER 1:

INTRODUCTION

1.1 TARGET DETECTION, TRACKING AND PURSUIT

Many insects have the capability to detect, track and chase moving targets either for prey or conspecifics (Land and Collett, 1974, Wagner, 1986a, Land, 1993). Although female houseflies will chase other flies in free flight (Wehrhahn, 1979) this behaviour is less accurate and of shorter duration than male pursuits. Male houseflies will chase another fly continuously, engaging in rapid and circling trajectories, resulting in high angular velocities of over $3000^\circ/\text{s}$ (Land and Collett, 1974). As would be expected, these behavioural differences have led to evolution of sexual dimorphism in the morphology and neurophysiology of the visual system (Dietrich, 1909, Collett and Land, 1975, Kirschfeld and Wenk, 1976). Specializations extend from the optical design of the eye, for example, the high acuity ‘love spot’ in the male housefly (Beersma et al., 1975, Land and Eckert, 1985), through to differences in the higher-order neurons (Hausen and Strausfeld, 1980, Strausfeld, 1991, O'Carroll et al., 1997, Nordström et al., 2008).

Male blowflies track their targets and control both forward velocity (dependent on angular size of the target) and turning velocity (dependent on angular position of the target) during chases (Boeddeker et al., 2003). In an entirely different method of chasing, insects such as large hoverflies (*Eristalis* and *Volucella*) and dragonflies will intercept their targets (Collett and Land, 1978, Olberg et al., 2000). These insects appear to derive sufficient information of target motion in early stages of the detection so as to plot an interception course.

In some species, a chasing fly will attempt to maintain the target in the dorso-frontal visual field (Wehrhahn et al., 1982). Whilst others have formed specialised eyes, with dorsal acute zones, to silhouette a target against the bright sky (Zeil, 1983). Still others have frontal acute zones (Land and Eckert, 1985) and therefore must track targets against complex visual textures. Dependent on the fly (species, gender and behaviour) some eyes show a structural form, tuned for the target pursuit task, both within the

distribution of facets (Land and Eckert, 1985), the dynamics of the photoreceptors (Hardie et al., 1981, Hornstein et al., 2000) and the other underlying neural architecture (Collett and King, 1975).

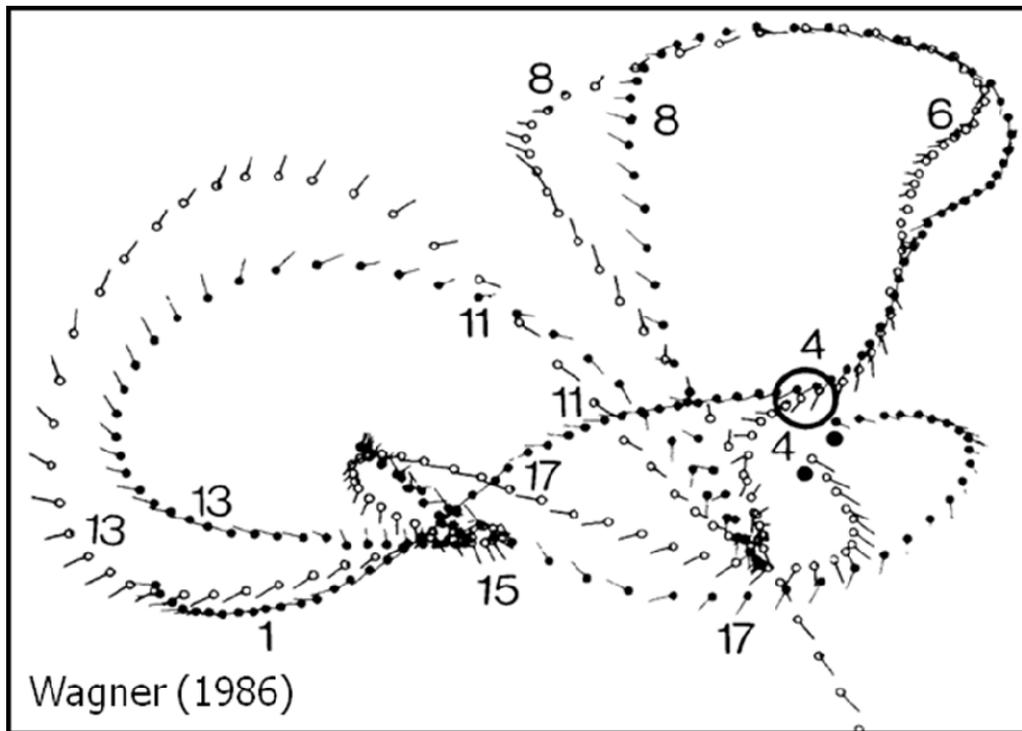


Figure 1-1. Fly chase behaviour. Reproduced from Wagner, 1986. Rapid turns at high angular velocities (over $3000^\circ/s$) typify flights, as flies pursue one another (point intervals 10ms).

In this thesis, I will investigate the neural pathways that may form the basis for detecting targets in visual clutter. I have also used electrophysiological recordings of neurons from various stages of the insect visual pathway, including from neurons that respond to small moving targets, to develop a biologically inspired model for this task.

1.2 MOTION DETECTION THEORY

During chase behaviour, the fly will encounter two general forms of motion. One is a rotatory and translatory movement of the entire visual field which will be induced by the fly's ego-motion (self motion). The other is the movement of objects within the visual surround. These require two separate motion detection systems, as was first observed in behavioural experiments with crickets (Palka, 1969), flies (Geiger and Poggio, 1975, Srinivasan and Bernard, 1977), and locusts (Rowell et al., 1977).

There are three primary methods to detect motion. The first involves identification and tracking of features, typically in frame-based representations of the moving scene, and is a popular method within the computer vision community. The other two types deal with 'first-order' visual motion, defined by a spatial displacement of luminance over time. The detection of these spatiotemporal luminance changes requires two spatially separated receiving elements, an asymmetry in the connected network and the nonlinear interaction of the two signals (Buchner, 1984). Two mathematically distinct families of models achieve this detection task, the gradient scheme (Limb and Murphy, 1975, Srinivasan, 1990) and motion energy or correlation schemes (Adelson and Bergen, 1985, Hassenstein and Reichardt, 1956). The three forms of motion computation are described in detail in the following sections.

1.2.1 Feature-tracking

At any instant of time, features (or feature derivatives) are identified within a scene and their spatial locations quantified (for examples, Harris and Stephens, 1988, Moravec, 1977, Shi and Tomasi, 1994, Smith and Brady, 1997). In the following frame (discrete time), the features are again spatially localised. The displacement of the features is calculated, and with the frame period known, the motion of the features or scene is readily calculated. The challenging part of this process is in identification of the location of features.

1.2.2 The Gradient Scheme

The 'gradient detector' is based on the computation of velocity by dividing the derivative of luminance over time, by the derivative of luminance over space. This method of motion detection is based on the following equation.

$$\frac{dx}{dt} = \frac{\frac{dI}{dt}}{\frac{dI}{dx}}$$

This calculation can be implemented with high-pass temporal filtering and a subtractive inhibitory interaction between two sampling points (an approximation). A divisive element is also required (Torre and Poggio, 1978) and the availability of such a neural element has been the subject of intense investigation (Holt and Koch, 1997). Interestingly, while there are numerous applications of models of this kind in artificial vision, evidence for neural mechanisms resembling gradient detection in biological systems is scarce, possible due to the difficulty of implementing the divisive operator.

The signal provided by the gradient detector is proportional to the image velocity, however, because of the division, if the spatial derivative is low then noise will be amplified (also, the signal is undefined for a spatial derivative value of zero). Elaborations to the gradient scheme have resolved these issues (Srinivasan, 1990, Sobey and Srinivasan, 1991) and this method of motion computation is commonly used within the image processing field.

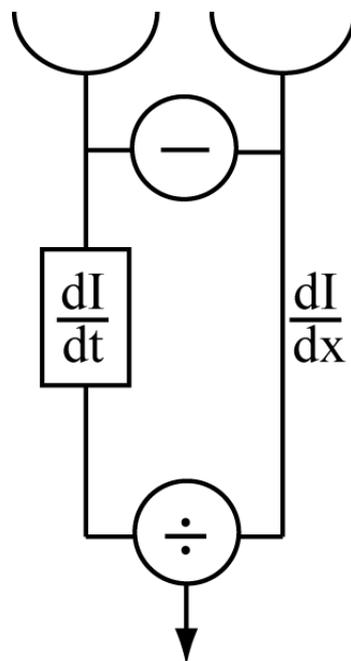


Figure 1-2. The Gradient Detector. Computation of velocity by dividing the derivative of luminance over time, by the derivative of luminance over space.

1.2.3 The Reichardt Correlator

In seminal studies, Hassenstein and Reichardt (1956) studied the behavioural turning response (the optomotor response) of the beetle *Chlorophanus* to the movement of the visual surround. This turning response is the animal's attempt to compensate for a perception of ego-motion in the opposite direction. The analysis of these results led to the development of a motion detection model referred to as a 'correlation-type detector' or 'Reichardt correlator'.

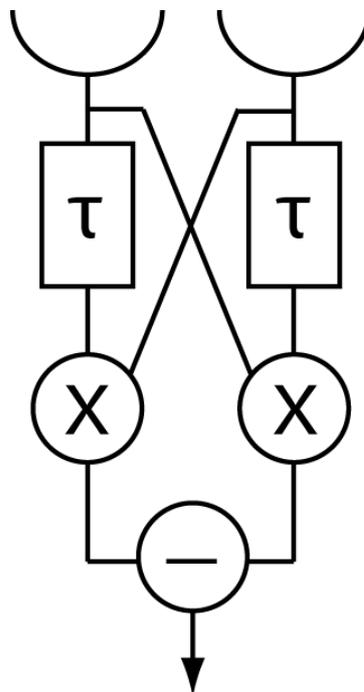


Figure 1-3. The Reichardt Correlator. A delay (τ) and compare (multiplication, x) of two spatially separated inputs signals the motion.

The correlator subunit consists of two spatially separated inputs, with one of the signals delayed relative to the other (e.g. via a low pass filter), before multiplication between the two arms (Hassenstein and Reichardt, 1956). More usually, the signals are initially processed via spatial and temporal prefilters (Reichardt, 1961, van Santen and Sperling, 1985). The prefilters represent early visual processing, such as the optical blurring, photoreceptor responsiveness and spatiotemporal filtering characteristics of 1st-order interneurons. The multiplication results in a signal that will (on average) be more positive when a pattern moves in the preferred direction (and more negative in the anti-

preferred direction), however, the unit is still very responsive to a flicker (non-motion) luminance component. Therefore, a mirror symmetric unit is subtracted from the original forming a full correlator mechanism. The Reichardt correlator is commonly referred to as an elementary motion detector (EMD).

Other forms of motion detection, such as the Energy model (Adelson and Bergen, 1985), have been found to be holistically equivalent to the correlation based mechanism, only with a different underlying mathematical architecture (Adelson and Bergen, 1985, van Santen and Sperling, 1985). The Reichardt correlator suggests a multiplicative unit for the comparison stage, though several other possibilities have also been suggested, for example, asymmetric and nonlinear inhibitory mechanisms in the rabbit retina (Barlow and Levick, 1965), turtle retina (Ariel and Adolph, 1985) and in the medulla of the fly (Mimura, 1972) and locust (Osorio, 1986). These proposed models, based on delayed inhibition, also have a similarity to the correlator-type detectors, as the ‘and-not’ structure approximates the multiplicative nonlinearity in a digital sense (Buchner, 1984). Further elaborated inhibitory models have provided explanation for facilitatory responses observed in the insect visual system (Bouzerdoux and Pinter, 1990, Bouzerdoux, 1993).

These ‘equivalent type’ detectors have been used to describe motion computation in a variety of biological contexts; in insects (Hassenstein and Reichardt, 1956), rabbit (Barlow and Levick, 1965), humans (van Santen and Sperling, 1983), wallabies (Ibbotson et al., 1994) and pigeons (Wolf-Oberhollenzer and Kirschfeld, 1994). The correlator has also explained behavioural experiments on the fly optomotor response (Fermi and Reichardt, 1963, Buchner, 1976). Hence, this form of motion detector is likely to have evolved numerous times and appears to be the ubiquitous biological motion detector.

The neural mechanisms that could implement the mathematical operation of multiplication are still yet to be determined, however plausible biophysical mechanisms have previously been suggested (Koch, 2004). The location in the visual pathway of the correlation mechanism, particularly the multiplicative processing, remains unknown.

By increasing the spatial phase (by changing the spatial sampling baseline of the inputs), or decreasing the time constant of the delay, the correlator-detector is tuned for higher velocities. An array of EMDs may be combined via spatial integration (either

linearly, or nonlinearly) to provide an element that signals directionally opponent wide-field motion. Higher order neurons within the optic lobes of the fly, the lobula plate tangential cells (LPTC), are found to have these characteristics (for example, Single and Borst, 1998, Haag et al., 2004).

However, this ‘tuning’ of the correlation model leads to interesting characteristics in coding velocity by these types of motion detector. The response of the direction selective neurons would not represent velocity in a linear manner, rather would increase to an ‘optimal velocity’ and then decrease as this velocity was exceeded (Buchner, 1984). This optimal velocity should also vary proportionally to the spatial frequency of the stimulus, maintaining a constant ratio between the two. Some of these effects have been verified experimentally (review Borst and Egelhaaf, 1989). However, many modifications of the EMD circuit have been proposed to more closely match the response characteristics of the recorded neurons. EMDs that are independent of the spatial structure of the image, including contrast invariance, are required for a complete model (Zanker et al., 1999).

For these reasons, many engineers have rejected the Reichardt correlator as ambiguous and unusable in attempts to build artificial optic flow detectors. This begs the question as to why this type of scheme has become so ubiquitous in nature.

It has been show that an ‘optimal’ motion detector, with respect to the signal-to-noise ratio (SNR), would be based on the gradient scheme under high SNR conditions, however under low SNR, a correlation mechanism is most appropriate (Potters and Bialek, 1994). However, Haag et al. (2004) showed that motion sensitive cells in the fly were based on Reichardt correlators over a large SNR range, i.e. over a wide range of luminance and contrast conditions.

1.2.4 The Motion Stimulus (Classical and Natural Images)

Having discussed motion detection theory, it is relevant to address what is actually moving. Traditional investigation of motion processing examined responses to ‘simple’ stimuli such as moving objects (dots and bars) and drifting (sinusoidal) gratings. Gratings have the benefit of a simple mathematical representation and analysis of input/output transforms that are more readily defined. Felson and Dan (2005) examine

arguments that biological motion is best examined with ‘natural’ inputs, whilst Rust and Movshon (2005) praise classical visual stimuli.

With respect to studying natural stimuli and vision processing there are two general approaches (Simoncelli and Olshausen, 2001). One method is to examine neural responses under natural stimulation conditions (Laughlin, 1981a, Rieke et al., 1995) whilst the other is to predict a visual processing model for a particular statistical optimization (relevant to natural inputs). This model is then compared with the neural responses (Atick, 1992, Olshausen and Field, 1996, van Hateren, 1998).

Natural images are found to have consistent spatial statistics. Their spectral power falls with increasing frequency (f), in a power law relationship $1/f^p$, ($p \sim 2$) (Field, 1987, Ruderman, 1994). There is a correlational structure within natural scenes and Srinivasan et al. (1982) showed that the amount of spatiotemporal antagonism (within a visual neuron of the fly) was matched to remove these correlations from typical moving natural scenes. Further predictions were made regarding the dynamic changes of these filtering characteristics due to information theoretic considerations (low-pass in low SNR, high-pass in high SNR), again dependent on the natural stimulus inputs (van Hateren, 1992c).

The Reichardt detector provides an ambiguous velocity signal (for differing combinations of spatial and temporal frequencies). However, it was shown that due to the predictable statistics of natural images, the mean response of spatially integrated EMDs does provide a reliable estimate of velocity (Dror et al., 2000). Residual problems remain in providing an accurate velocity estimation which would match the robust velocity responses observed behaviourally in the bee (Srinivasan et al., 1991). The study of spatially integrated, wide-field EMDs that can provide a robust estimation of velocity is presently a focus of research (Shoemaker et al., 2005, Brinkworth and O'Carroll, 2007)

1.3 INSECT VISUAL PATHWAY

This thesis is concerned with a variety of insect visual systems and aims at generalising target discrimination across the visual systems of blowflies, hoverflies and dragonflies. In this brief review of insect visual pathways, I will focus on the fly visual system which provides the required background information. However, there are significant differences between the insect species, and where relevant these are highlighted.

1.3.1 Optics

Light intensity of the surrounding environment is processed through the optics of the compound eye via thousands of (semi) regularly arranged facet lenses. Each facet points in a different direction and accepts light from a narrow angle (review Nilsson, 1989, Strausfeld, 1989). Light is guided to eight photoreceptors contained within a single sampling unit referred to as an ommatidium. The optics of male houseflies is tuned to the task of target detection. In an ‘acute zone’ they have larger ommatidial facet lenses (Land and Eckert, 1985) which have higher acuity and better sensitivity (Land, 1981).

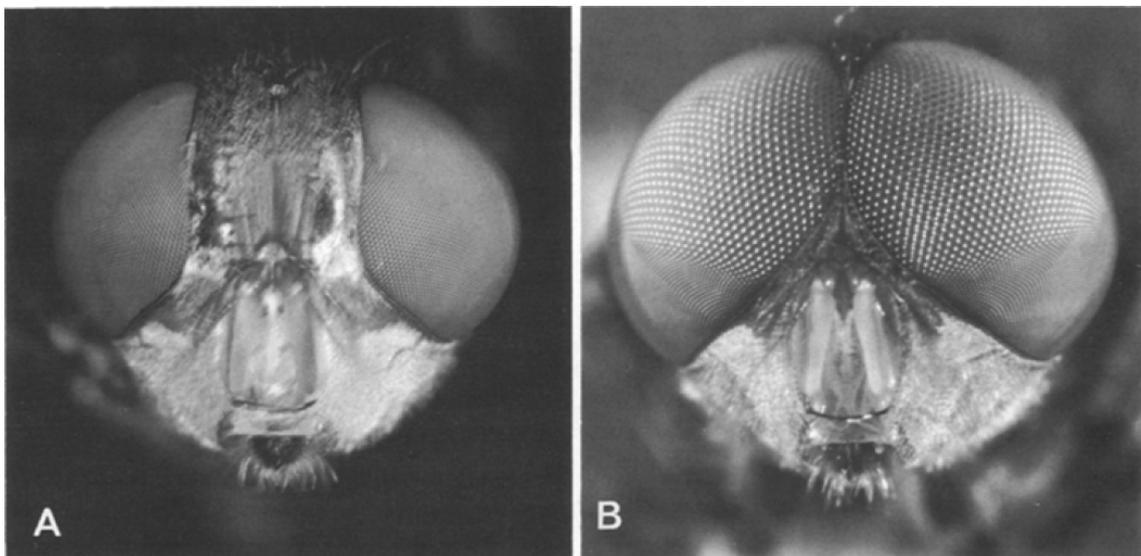


Figure 1-4. Sexual Dimorphism. Reproduced from van Hateren et al (1989). (A) Female and (B) Male *Chrysomyia*. The dorsal eye of the male has greatly enlarged facets. This ‘bright’ zone, rather than an ‘acute’ zone, provides greater sensitivity in this region of the eye.

1.3.2 Photoreceptors

The photoreceptor cells act as light sensors, producing graded membrane potentials via the process of phototransduction (review Yarfitz and Hurley, 1994, Hardie, 2001), and then transferring this information to the visual interneurons in the lamina.

Adaptive mechanisms allow the photoreceptors to operate in a large range of luminance conditions (Laughlin, 1994), and shift their behaviour dynamically dependent on changing background luminance. Photoreceptors code for contrast (relative differences in intensity, both in space and time) (Barlow, 1961, Laughlin et al., 1987) and alter their contrast gain to reduce noise and improve the SNR dependent on the adaptation conditions (Juusola et al., 1994). Generally, photoreceptors act as low-pass filters that change to band-pass filters with increasing light adaptation (Jarvilehto and Zettler, 1971).

Photoreceptors have to encode the large dynamic range encountered in natural scenes, into the limited dynamic range of the neuron (van Hateren, 1992c). Dependent on the spatial power spectrum of the scene and the behaviour of the insect (velocity dependent temporal structure) theoretically optimal filters have been developed that maximise information transmission (van Hateren, 1992a, van Hateren, 1993). A parametric photoreceptor model has been developed, based on a cascade of linear and nonlinear elements that matches information theoretic expectations of photoreceptor responses to natural scenes (van Hateren and Snippe, 2001).

Interestingly, the complex and dynamical properties of the photoreceptor has recently been found to enhance target salience within moving scenes, purely by their temporal processing characteristics (Brinkworth et al., 2008).

1.3.3 Neural Superposition

Previous research has provided detailed description of the neural interconnections within the retina (Strausfeld and Nässel, 1981). Six photoreceptors (R1-R6), from adjacent ommatidia, looking at the same point in space, converge onto a single lamina cartridge by the principles of neural superposition (Nilsson, 1989). The inner photoreceptors (R7 and R8) are involved in colour vision, bypassing the lamina and terminating in the medulla (Hardie, 1985). Recent experiments in *Drosophila* (Yamaguchi et al., 2008) have verified that only the outer photoreceptors (R1 to R6) are

involved in the motion processing pathway. These photoreceptors are sensitive to UV or green-blue light (review Hardie, 1985).

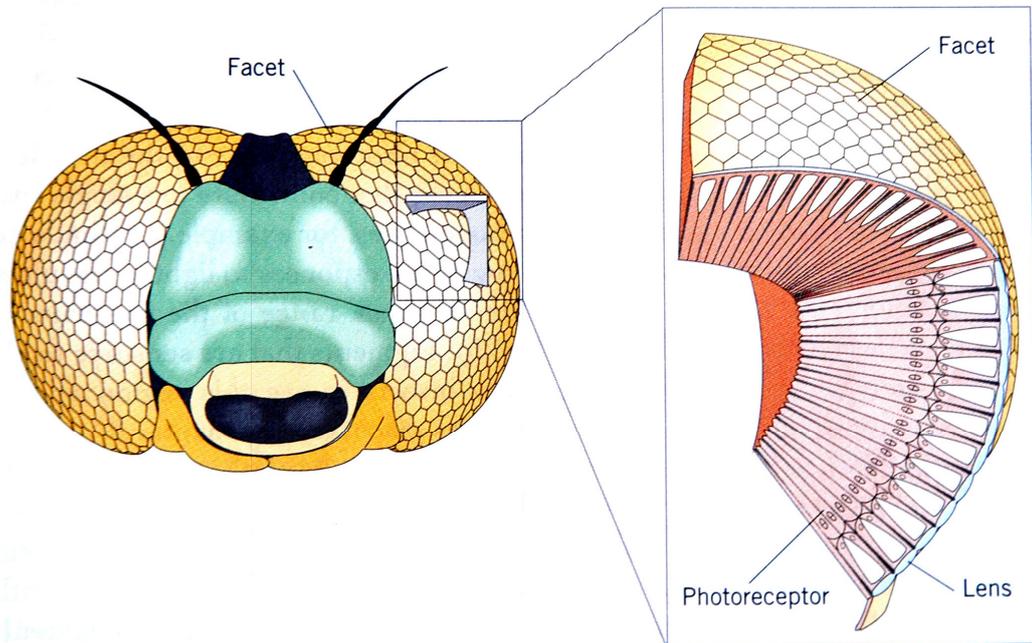


Figure 1-5. Fly Eyes. Reproduced from Delcomyn (1998). This diagram shows the arrangement of the facets and the underlying structure of the lens which guide light towards the photoreceptors.

1.3.4 Lamina Neurons

The 2nd optic ganglion, the lamina, is usually regarded as the principal site of redundancy reduction in the insect visual pathway, with clear evidence for both spatial and temporal high-pass filtering in the Large Monopolar Cells (LMC) (Laughlin et al., 1987).

Neurons within the lamina neuropile are grouped into cartridges, referred to as neuro-ommatidia, and each grouping corresponds to one sampling point of space. This columnar (retinotopic) organisation is maintained through the visual stages. In the lamina, the visual signal is processed by three large monopolar cells (LMC) (L1, L2 and L3), two small monopolar cells (SMC) (L4 and L5), an amacrine cell (AM) and the basket cell T1 (Laughlin, 1984, Shaw, 1984).

The graded responses of the three LMCs are generally the same, with an ‘on’ hyperpolarization, a sustaining component and a transient ‘off’ depolarization. These

components vary depending on stimulus duration and the state of light adaptation (Laughlin, 1981b). They are not motion specific (Gilbert et al., 1991). There are small physiological and anatomical differences between L1 and L3 cells compared with L2 neurons, which may suggest a slight difference in the temporal dynamics for these neuron types (Hardie and Weckström, 1990). Histamine neurotransmitter is being tonically released by the photoreceptor, even in darkness (Laughlin et al., 1987), and an increase in release causes histamine gated chloride channels to open, hyperpolarizing the LMC (Hardie, 1987). This metabolically expensive processing allows for the rapid signalling of increased and decreased contrast changes.

The LMC, due to spatial and temporal antagonism, removes redundant information from correlations in visual scenes (Srinivasan et al., 1982). A more thorough examination reveals that the LMC alters its filtering characteristics dependent on visual conditions. In the dark adapted 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 (Juusola et al., 1995). These transient responses enhance edge-like boundaries, to varying degrees depending on adaptation state (Srinivasan et al., 1990b), which enhances target discrimination (Wiederman et al., 2008). Spatial antagonism observed in the LMCs appear to utilize inhibitory interactions between nearest neighbour receptors (Srinivasan et al., 1982).

The less well known L5 is a spiking neuron, which seems to replicate photoreceptor dynamics with a sustained response to light and adaptation characteristics (Jansonius and Van Hateren, 1993b). The basket T1 neuron has dendrites that receive information from R1-R6 via the amacrine cell (Campos-Ortega and Strausfeld, 1973). The medulla terminals of T1 mix with each L2 terminal and the dendrites of the transmedullary cell, Tm1. Tm1 has a non-directional movement sensitive nature and has been hypothesised to play a role in the motion detection pathway (Gilbert et al., 1991, Douglass and Strausfeld, 1995).

There is still a vast amount of understanding to be developed regarding the functionality of the insect lamina neuropile. Defined roles for L4 and L5 remain unknown and it is possible that the amacrine cell may not only play a part in lateral inhibition (Strausfeld and Campos-Ortega, 1977), but may also be intrinsic to the motion detector delay

pathway. Differentiating roles for the large monopolar cells (L1, L2, L3) are unknown, though responses are well characterised. Electrophysiological responses, presumed to be from the L4 neurons, are discussed in the following section. In the coming years, with new genetic intervention techniques and further electrophysiological study, it is hoped that some of these unknowns may be addressed (Keller, 2002).

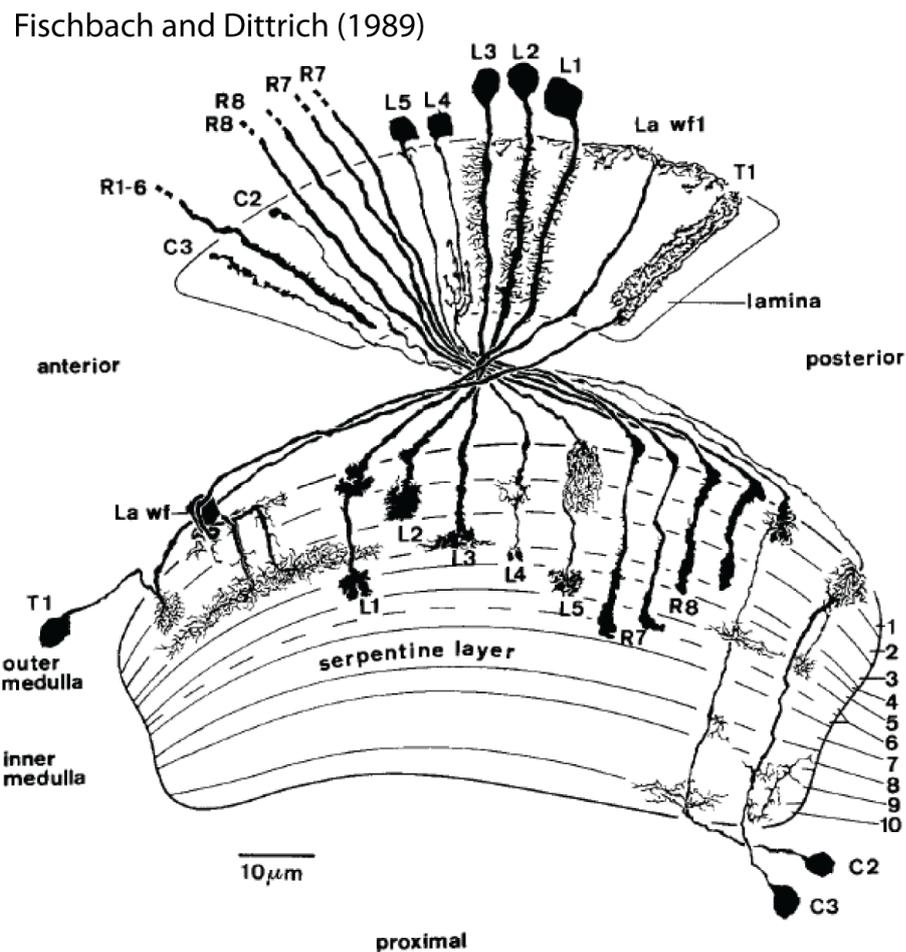


Figure 1-6. The Lamina and Medulla of *Drosophila*. Reproduced from Fischbach and Dittrich (1989). Photoreceptors, small and large monopolar cells and other neuron types compose the elements of early visual processing.

1.3.5 On-Off cells

Extracellular recordings in the first optic chiasm between 2nd and 3rd order interneurons of the fly brain (between the lamina and medulla), first showed the presence of ‘on-off’ cells (Arnett fibers) with full-wave rectification (Arnett, 1971, Arnett, 1972). This work included lesion experiments that showed these fibres to be centripetal, indicating an impressive level of complexity at an early processing stage of the visual pathway (i.e., the lamina). These cells (presumed to be a SMC) were later re-examined and shown to adapt independently to luminance changes, dependent on the polarity (increment or decrement) of the change (Jansonius and van Hateren, 1991). Contrast sensitivity to wide-field stimulation was found to be relatively poor. Experiments with moving sinusoidal gratings indicated a low-pass temporal characteristic with roll-off of response at the relatively low frequency of ~12 Hz, compared with the much higher temporal responsiveness of the photoreceptors (Jansonius and van Hateren, 1991). Independent adaptation has also been seen in locust medullary cells (Osorio, 1991) which was hypothesised to detect contrasting edges of visual features, ignoring their inner “textual detail” (Osorio, 1987, O’Carroll et al., 1992).

To examine spatial interactions on a larger spatial scale, visual LED stimuli of separated pulses (5°) were presented to the ‘on-off’ cells (Jansonius and Van Hateren, 1993a). This revealed spatial antagonism over a distance equivalent to several facets of the compound eye.

Neural models have been developed for these transient cells and were proposed to serve a role in the front end of motion detection (Ogmen and Gagne, 1990). The transient cell model of Sarikaya et al. (1998) in particular is biophysically detailed, and accounts for many of the temporal characteristics as observed in the electrophysiological studies in the fly.

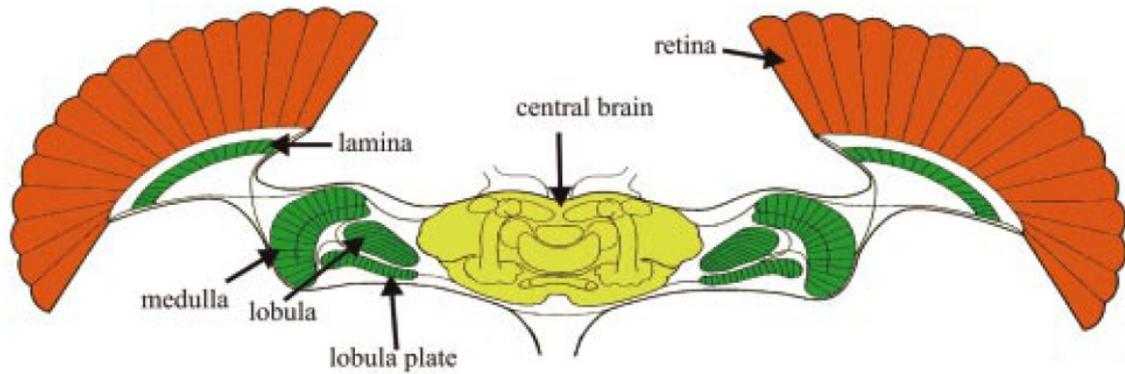


Figure 1-7. Cross-section of the Fly Head. Reproduced from Borst and Haag (2002). This image shows the visual processing regions, the retina, lamina, medulla, lobula and lobula plate.

1.3.6 The Neural Basis for an Elementary Motion Detector (EMD)

Presently, the scientific consensus is that there are no directionally selective neurons in the lamina of the fly (Mimura, 1972), however, evidence has been provided for phase shifts in the response of a lamina neuron, dependent on the direction of motion (Douglass and Strausfeld, 1995). Motion-sensitive cells that arise in the medulla have been identified via activity (motion) dependent staining (Buchner, 1984, Bausenwein and Fischbach, 1992). Due to their small size, medullary neurons are difficult to record from using electrophysiological techniques. Therefore, determining the components of the EMD, which may occur in this neuropil, is a challenging task. There is a proposed motion pathway, building on non-directional movement sensitive cells (responsive to movement in any direction) and passing through the T5 cells of the lobula (Gilbert et al., 1991). These studies have developed over many years (Douglass and Strausfeld, 1995, Douglass and Strausfeld, 1996, Douglass and Strausfeld, 2003) and have served as the basis for computational motion models which correspond to the electrophysiological findings (Higgins et al., 2004).

An alternative hypothesis is that EMDs are contained within the dendritic processing of the higher-order LPTC neurons (Single et al., 1997) which spatially integrate motion signals (see 1.4.1).

The initial stage of motion detection models are often designed with lamina LMC neurons delivering high-pass filtered information to the correlator inputs. An earlier study outlined reasons that LMC's were not on the motion pathway (Coombe et al., 1989). However, the LMCs are ideally suited to remove redundancy (Srinivasan et al.,

1982), maximize information transmission (van Hateren, 1992b) and due to their high-pass filtering nature (suitable for modelled EMDs) should be considered as plausible neural elements in the motion detection pathway.

There have been limited successful recordings from neurons within the medullary neuropil. A variety of response types have been reported (DeVoe and Ockleford, 1976), including reference to ON/OFF transients, sustaining neurons, spiking and non-spiking varieties of cells, those with directional selectivity, and those sensitive to motion in any direction. In research on the locust, further and more complete investigation of the medullary neurons has been completed (Osorio, 1986, Osorio, 1987).

1.4 NEURAL MECHANISMS AND MODELS FOR MOTION DETECTION

In this section I explore several unique classes of motion detecting neurons described from insects. In each case, these neurons have played an important role in the development of neuromorphic models for motion detection.

1.4.1 Lobula Plate Tangential Cells (LPTC)

Electrophysiological recordings from wide-field motion-sensitive cells in the lobula plate (a higher-order visual region of the fly brain) indicate an underlying motion framework based on the Reichardt correlator (review Borst and Egelhaaf, 1989). These 30-60 Lobula Plate Tangential Cells (LPTCs) (Hausen, 1982), are sensitive to visual motion in a directional way (review Borst and Haag, 2002). The cells' optimal response are matched to directions of rotation around the flies' own axis (yaw, pitch and roll) (Krapp and Hengstenberg, 1996). The LPTCs have been shown to spatially integrate over arrays of EMDs (Franceschini et al., 1989).

1.4.1.1 Figure Detection (FD) neurons

A subset of LPTCs, referred to as the FD neurons, have been shown to respond to medium size (rather than wide-field) motion (Egelhaaf, 1985b). Originally, a series of experiments that showed fly fixation behaviour, with discrimination between the figure and ground of moving random-dot patterns (Reichardt and Poggio, 1979), led to the development of a model for the visual detection of objects moving relative to backgrounds (see 1.5.2) and the postulation of the existence of a corresponding cell type (Reichardt et al., 1983). As modelled, these figure-detection (FD) cells are most sensitive to the movement of smaller objects, though not to wide-field motion of the visual surround. These cells were later discovered in the lobula plate of the fly optic lobes using electrophysiological techniques (Egelhaaf, 1985b).

Evidence for relative motion calculations to underlie the FD neural response was observed in ablation experiments in the fly. When the CH neuron, an LPTC, was laser-ablated, object specificity was removed (Warzecha et al., 1993, Egelhaaf and Borst, 1993). However, later experiments revealed no strong dependence of FD cell responses to the velocity of the background motion (on which CH cell response depends strongly) (Kimmerle and Egelhaaf, 2000). This suggests the relationship between object

selectivity and possible inhibitory effects from background motion may be more complex. Higgins and Pant (2004) developed and simulated an elaboration of the FD model which provided an explanation for this unexpected result via more complex interaction between the wide-field neurons (see 1.5.3).

FD neurons have a peak response to stimuli (composed of limited spatial extent gratings) of greater than 10° in angular size. As the grating stimuli size increases further, neural responses roll off relatively slowly. FD neurons are motion opponent (Egelhaaf, 1985b) expressing strong directionality (excited by motion in the preferred direction, whilst inhibited from motion in the anti-preferred direction). This suggests that like the wide-field LPTCs, FD neurons are built from a Reichardt correlator framework.

1.4.2 LGMD/DCMD

The lobula giant motion detector (LGMD) is a neuron located within the third optic neuropil of the locust and is sensitive to a ‘looming’ stimulus (Strausfeld and Nüssel, 1980). The neuron has three large dendritic regions, the largest receiving excitatory local motion inputs, with the other two receiving feed-forward, size dependent inhibition (Rowell et al., 1977, O’Shea and Williams, 1974). The LGMD synapses onto the descending contralateral motion detector (DCMD), in a one-to-one spiking relationship (O’Shea and Williams, 1974). The DCMD projects to motor centres responsible for jump and flight steering (Pearson et al., 1980) and has been suggested to be responsible for escape and collision avoidance behaviour when the looming visual stimulus reaches a fixed angular threshold size (Hatsopoulos et al., 1995).

Before the looming nature of the LGMD/DCMD system was determined (Rind and Simmons, 1992), the neural system was thought to be selective for small targets (Rowell et al., 1977). This size selectivity was modeled by lateral inhibitory interactions around a centre unit. This modelling was based on cells within the locust medulla responding transiently to both contrast increments (‘on’ channel) and contrast decrements (‘off’ channel) in a full-wave rectified manner (see 1.3.5). A lateral unit, derived from the local signal spread of these channels, was identified in the proximal medulla (O’Carroll et al., 1992). These were hypothesized to mediate the inhibitory interactions on the centre units and are still likely to play a role in the looming response of the LGMD.

1.4.3 Target selective neurons in the insect brain

Within the insect brain there are a variety of neurons that have smaller size selectivity than that of the FD neurons. These cell classes may underlie pursuit behaviour as they are more appropriately size tuned to the moving targets. Neurons suitable for this task have been observed in the hawkmoth (Collett, 1971), hoverfly (Collett and King, 1975) and dragonfly (Olberg, 1981, Olberg, 1986, O'Carroll, 1993).

1.4.3.1 Small Target Motion Detectors (STMD)

New characterizations of target-selective neurons, within the optic ganglia of the hoverfly and dragonfly, have recently been investigated (Nordström and O'Carroll, 2006, Nordström et al., 2006, Barnett et al., 2007, Geurten et al., 2007). Unlike the relative 'larger object' selectivity of the FD neurons, these 'small target motion detectors' (STMDs) are found to be selective for small targets of no more than a few degrees of the visual field. The receptive fields of STMDs are varied in size, i.e. the area in which the visual presentation of a small target elicits an excitatory response. Some receptive fields (the small-field STMDs) extend just a few degrees (O'Carroll, 1993) and are retinotopically organised (Barnett et al., 2007), whilst others cover a large part of the visual field. The target response may vary in magnitude across this region, however the size selectivity is independent of the target location or the size and shape of the receptive field (Nordström et al., 2006).

The STMD neural response varies with the changing of target parameters, expressing an optimal target size and velocity tuning, whilst also exhibiting a strengthening of response to increased target contrast.

Some STMDs respond to targets moving relative to a background, in many cases when the background itself is moving (Nordström et al., 2006). As with the FD neuron, it would seem likely that segregation of the motion of the target from the motion of the background would be a possible mechanism for the target discrimination. However, a subset of the larger-field STMDs respond robustly even when the velocity of the target is matched to the velocity of the background, i.e. with no relative motion cues (Nordström et al., 2006). If only wide-field background motion is presented to the STMD, the cell does not respond. This implies that the spatial statistics of small targets, with respect to the background, form an important cue for discrimination, regardless of any additional role that may be played by other motion cues (Nordström et al., 2006).

Within this thesis I present a model for target discrimination in visual clutter, which is derived from a proposed neural pathway to the STMD. This model relies on the spatiotemporal statistics of the moving target rather than relative motion cues between target and background.

1.5 NEUROMORPHIC MODELS FOR TARGET SELECTIVITY

1.5.1 Photoreceptors

The target detection model, as described in this thesis (see 4.2), represents early visual processing mechanisms by including an elaborated version of a photoreceptor model first devised by van Hateren and Snippe (2001). This model of photoreceptor dynamics was determined by varying parameters in a suitable cascade of linear and nonlinear elements (with feedbacks) until the coherence between stimulus and response best matched the rate seen within a physiological context. This model consisted of adaptive temporal filtering, followed by two divisive feedbacks (at different speeds – slow and fast adaptation) and a saturating nonlinearity. Brinkworth et al. (2008) perform electrophysiological recordings from photoreceptors that were visually stimulated with natural time series images (van Hateren, 1997). The natural series contained embedded targets and the photoreceptor processing revealed enhanced target salience, even without any spatial interactions. This was due to the temporal nonlinear dynamics of the photoreceptor.

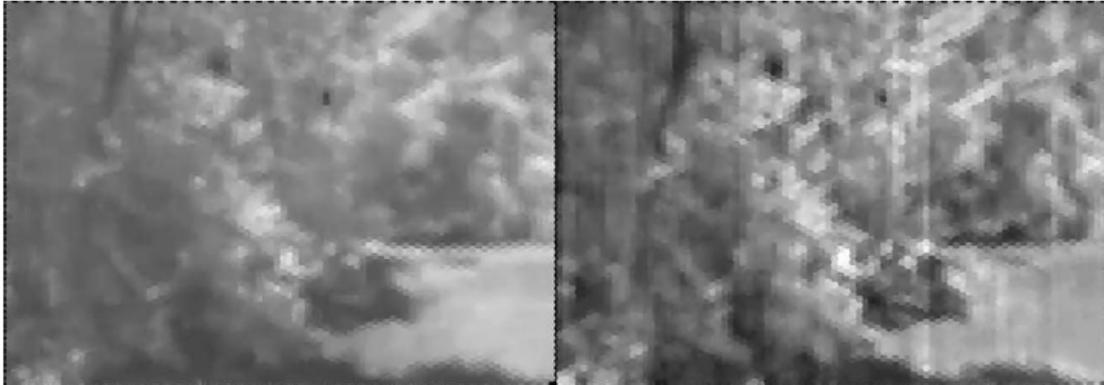


Figure 1-8. Photoreceptor Processing. Reproduced from Brinkworth et al. (2008). Left is the original luminance image and right is the reconstruction from electrophysiological recordings of the photoreceptor, showing an enhancement of target salience.

1.5.2 Figure Detection (FD) Cells

Models of FD cells vary dependent on the particular FD type (Egelhaaf, 1985a). Generally speaking, EMDs are summed to form directional (or non-directional) motion ‘pool’ cells. These progressive or regressive motion units are coupled across the eyes so as to provide representations of rotational motion. The pool cells use shunting inhibition

1.5.4 STMD

To date there has been limited modelling of the STMD neuron. In Guerten et al (2007), dragonfly centrifugal small target motion detector (CSTMD) neurons were stimulated with varying width/velocity profiles of small moving targets. These responses suggested an EMD type mechanism underlying the computation of target motion.

In Chapter 2 and 3 of this thesis, I develop a model for small target motion detection inspired by the properties of the STMD. This model includes an ‘on-off’ type component, based on my electrophysiological recordings from rectifying transient cells in the blowfly medulla.

In Chapter 4 and 5, I present an extended version of the model, including complex early visual processing mechanisms. I test the target discrimination mechanism under real world luminance and contrast conditions and across more general stimulus conditions of varying target sizes and velocities. Elaborations to the model allow for an optimal static threshold for the target discrimination under varying ego-motion conditions.

In Chapter 6, I test whether subtle manipulation of the visual stimulus can induce changes in the electrophysiological responses from STMD neurons in the dragonfly, as predicted by the model.

CHAPTER 2:

RECTIFYING TRANSIENT CELLS AND TARGET DISCRIMINATION

2.1 CONTEXT

Previous literature had speculated on roles for the ‘on-off’ units (see 1.3.5), for example, as inputs to motion correlators (Ogmen and Gagne, 1988), moving object boundary detection (Osorio, 1991) and as a modulator of the temporal behaviour of motion detectors (Jansonius and Van Hateren, 1993a).

Conceptually, this form of nonlinear processing (independent contrast polarity adaptation) would be suitable as a target discrimination mechanism. While the inherent properties of such cells are interesting for modelling, the limitation of their temporal responsiveness (Jansonius and van Hateren, 1991) seemed to preclude their role in target discrimination.

To clarify this issue we performed intracellular recordings from ‘on-off’ units in the medulla of the blowfly. This allowed us to re-examine the temporal responsiveness of a type of ‘on-off’ unit which we now refer to as the Rectifying Transient Cell (RTC). We found that taking into account their highly adaptive nature was an important aspect when examining their responses to visual stimuli. That is, when examining transient phenomena, the steady-state responses have little functional meaning. When we examined the transient, unadapted response of the RTC we found that these ‘on-off’ units were as fast as neurons found earlier in the visual system. A role in target detection was therefore feasible.

In this chapter we outline these physiological experiments and also test a model (which includes an RTC element) for the ability to discriminate targets from a series of natural images.

Due to the breadth of the publication, we limited our simulations to a series of four 8-bit natural images, with embedded targets of a single size and a single panorama velocity. Although limited, these simulations revealed that this form of nonlinear processing had, as expected, the ability to greatly aid in the discrimination of targets from within natural scenes.

Importantly, the physiological characteristics of actual STMD neurons, such as contrast sensitivity, target height tuning and velocity tuning, were also matched by the model output.

2.2 A MODEL FOR THE DETECTION OF MOVING TARGETS IN VISUAL CLUTTER INSPIRED BY INSECT PHYSIOLOGY

S. D. WIEDERMAN¹, P.A. SHOEMAKER², D.C. O'CARROLL¹

¹*Discipline of Physiology, School of Molecular and Biomedical Science, The University of Adelaide, Adelaide, Australia*

²*Tanner Research Inc., Monrovia, California, USA*

PUBLISHED: Wiederman S.D., Shoemaker P.A., O'Carroll D.C. (2008) A model for the detection of moving targets in visual clutter inspired by insect physiology. PLoS ONE 3(7): e2784 doi:10.1371/journal.pone.0002784

2.2.1 Abstract

We present a computational model for target discrimination based on intracellular recordings from neurons in the fly visual system. Determining how insects detect and track small moving features, often against cluttered moving backgrounds, is an intriguing challenge, both from a physiological and a computational perspective. Previous research has characterized higher-order neurons within the fly brain, known as 'small target motion detectors' (STMD), that respond robustly to moving features, even when the velocity of the target is matched to the background (i.e. with no relative motion cues). We recorded from intermediate-order neurons in the fly visual system that are well suited as a component along the target detection pathway. This full-wave rectifying, transient cell (RTC) reveals independent adaptation to luminance changes of opposite signs (suggesting separate ON and OFF channels) and fast adaptive temporal mechanisms, similar to other cell types previously described. From this physiological data we have created a numerical model for target discrimination. This model includes nonlinear filtering based on the fly optics, the photoreceptors, the 1st order interneurons (Large Monopolar Cells), and the newly derived parameters for the RTC. We show that our RTC-based target detection model is well matched to properties described for the STMDs, such as contrast sensitivity, height tuning and velocity tuning. The model output shows that the spatiotemporal profile of small targets is sufficiently rare within

natural scene imagery to allow our highly nonlinear ‘matched filter’ to successfully detect most targets from the background. Importantly, this model can explain this type of feature discrimination without the need for relative motion cues.

2.2.2 Introduction

Certain flies (as well as other kinds of insects) detect and track small moving objects as they engage in rapid pursuits, demonstrating the capability to discriminate between targets (e.g. other flies) and an often cluttered, moving background (Wagner, 1986b, Wehrhahn, 1979). This is an especially challenging task considering that the fly compound eye limits visual resolution to $\sim 1^\circ$ (Land, 1997).

Neurons sensitive to (and in some cases selective for) small moving targets have been described in a variety of insect species (Collett and King, 1975, Gilbert and Strausfeld, 1991, O'Carroll, 1993, Wachenfeld, 1994). Recent intracellular investigations have more carefully characterized a number of target-selective neurons in the optic ganglia of the hoverfly (Nordström and O'Carroll, 2006, Nordström et al., 2006, Barnett et al., 2007). These ‘small target motion detectors’ (STMDs) were found to be exquisitely selective for small targets subtending no more than a few degrees of the visual field, equivalent to just one or two ‘pixels’ of the compound eye. The receptive fields of STMDs vary in size, with some extending just a few degrees, to those that encompass the whole eye hemifield. The target response may vary in magnitude across this region, however the size selectivity is independent of the target location (Nordström and O'Carroll, 2006) or the size and shape of the receptive field (Nordström et al., 2006).

STMDs respond to targets moving relative to a background, in many cases when the background itself is moving (Nordström et al., 2006). Conceptually, it would seem likely that neural mechanisms required for such a task involve segregation of the motion of the target from the motion of the background. Surprisingly, whilst some STMDs exhibit a suppressed response in the presence of background motion, a subset respond robustly even when the targets move at the *same velocity* as the background, i.e. with no relative motion cues (Nordström et al., 2006). However, the response to wide-field background motion alone elicits no response. This implies that the spatial statistics of small targets, with respect to the background, form an important cue for discrimination, regardless of any additional role that may be played by other motion cues (Nordström et al., 2006).

2.2.2.1 Computational models for target discrimination

Understanding the computation that underlies small target selectivity and rejection of background motion presents a daunting challenge. Some models for target discrimination rely on inhibitory feedback of wide-field motion signals to localized motion detectors (Egelhaaf, 1985c, Higgins and Pant, 2004), which may provide an explanation for small target selectivity, but would lead to inhibition by background motion. Another model, for what some thought at the time was the target selectivity of a higher order locust neuron (Rowell et al., 1977), has lateral inhibitory interactions around a centre unit. This model was based on cells responding transiently to both contrast increments (ON channel) and contrast decrements (OFF channel) in a full-wave rectified manner. A lateral unit, derived from the local signal spread of these channels, was hypothesized to mediate the inhibitory interactions on these centre units (O'Carroll et al., 1992). Here we examine and model a similar neuron type we refer to as the 'Rectifying Transient Cell' (RTC). We show that fast temporal adaptation and lateral inhibitory connections, characteristic properties of RTCs, could provide the basis for an alternative model for small target selectivity, robust against wide-field background motion.

2.2.2.2 Full-wave rectifying transient neurons

Extracellular recordings in the first optic chiasm between 2nd and 3rd order interneurons of the fly brain (between the lamina and medulla), first showed the presence of "on-off" cells (Arnett fibers) with full-wave rectification (Arnett, 1971, Arnett, 1972). Surprisingly, these cells were later re-examined and shown to adapt independently to luminance changes, dependent on the polarity (increment or decrement) of the change (Jansonius and van Hateren, 1991). This independent adaptation was also observed in medullary neurons in the locust (Osorio, 1991). These locust neurons had a 'breakthrough response' when stimulated with a pulse of the same polarity but greater contrast than the prior adaptor. The authors hypothesized that this nonlinear adaptation might enhance responses to the contrasting edges of visual features, whilst rejecting lower contrast "textual detail" (Osorio, 1987, O'Carroll et al., 1992).

Spatial antagonism observed in the LMC, an earlier 1st order interneuron in the lamina, appears to utilize inhibitory interactions between nearest neighbor receptors (Srinivasan et al., 1982). However in the 'on-off' cell experiments (Jansonius and Van Hateren, 1993a), separated pulses (5°) revealed antagonism on a larger spatial scale, equivalent

to several facets of the compound eye. The authors proposed a model where rectification occurs after lateral inhibition of the subunits (Figure 1A), however, unless the inhibitory influence of neighbors is excessively strong, it is difficult to explain why the summing of spatially separated rectified signals, responding to pulses of ‘like’ sign, should produce an inhibition of the overall response as was observed (Jansonius and Van Hateren, 1993a). These results lead us to propose the possibility of a second-order of local inhibitory interactions between ‘like’ ON channels and OFF channels, before they are recombined via spatial pooling (Figure 1B).

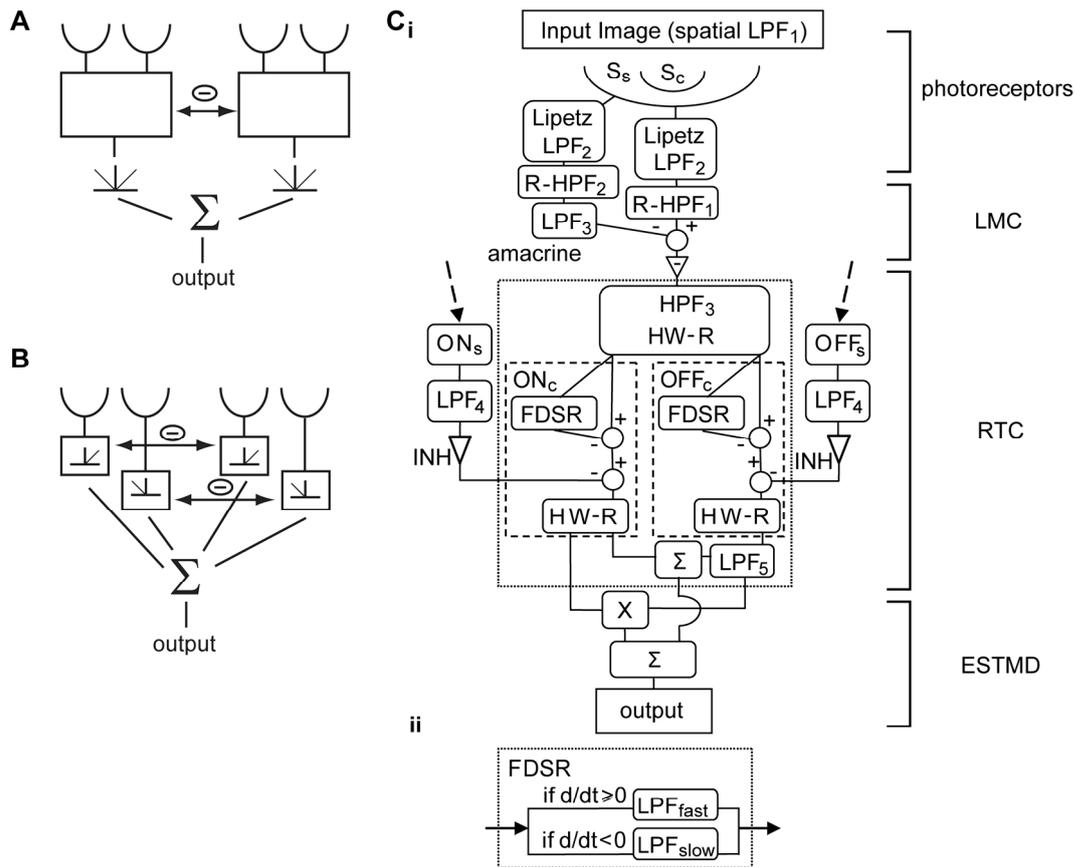


Figure 2-1. Model Overviews. (A) Model of the ‘on-off’ unit as described by Jansonius and van Hateren (Jansonius and Van Hateren, 1993a). The transient subunits exhibit fast adaptation and lateral inhibition before full-wave rectification and spatial pooling. (B) Our proposed version of rectifying transient cells where fast adaptive units are segregated into ON and OFF channels via half-wave rectification. Each polarity channel laterally inhibits one another before spatial pooling. (Ci) The detailed block diagram of the elementary small target motion detector (ESTMD) model. Early visual processing (photoreceptors, Large Monopolar Cells (LMC) and amacrine cells) is modeled with optical blurring (LPF₁), a nonlinear compressive transform (Lipetz function) with an adaptive mid-point parameter, and spatiotemporal band-pass filtering (LPF_{2&3}, R-HPF_{1&2}). The signal is separated into independent channels (responding to contrast increments and decrements respectively) via further high-pass filtering (HPF₃) and half wave rectification (HW-R). Each channel exhibits fast adaptation, implemented via the FDSR inhibition (see Cii). The channels are separately inhibited by a delayed (LPF₄) signal derived from surrounding channels of the same type. The strength of this surround inhibition is determined by the free gain parameter INH. To implement sensitivity to dark targets, the OFF channel is delayed (LPF₅) and then recombined with an undelayed ON channel in either a linear (Σ) or quadratic (X) manner. (Cii) Fast depolarization, slow repolarization (FDSR). If the input signal is ‘depolarizing’ (positive temporal gradient), a first-order low pass filter with a small time constant (LPF_{fast}) is used, otherwise for a ‘repolarizing’ signal (negative gradient) a larger time constant is applied (LPF_{slow}). The resulting processed signal represents an ‘adaptation state’ which then subtractively inhibits the unaltered pass-through signal.

[LPF₁ Gaussian blur (half-width 1.4°); LPF₂ $\tau=2.5$ ms; R-HPF_{1&2&3} $\tau=40$ ms, 30% DC; LPF_{3&4} $\tau=2$ ms; LPF₅ $\tau=25$ ms; LPF_{fast} $\tau=1$ ms; LPF_{slow} $\tau=100$ ms; INH = 3 (free parameter)]

2.2.2.3 Rectifying Transient Cells in the target detection pathway

We have developed a model for small target motion detection inspired by the properties of the higher order STMDs, and including a RTC-type component. We validate key stages of the model with intracellular recordings of the RTC in the fly (*Calliphora stygia*) medulla and with published physiological data. We investigate the temporal responsiveness of the RTC and obtain filtering parameters for the STMD model. We show that the properties of independent adaptation and lateral inhibitory interactions, as observed in ‘on-off’ cells and the RTC, are well suited for a role in target detection. We show that the spatiotemporal signature associated with the motion of a small feature is the passing of two contrast boundaries of opposite polarities (i.e. due to the leading and trailing edges), with limited spatial extent – which induces an excitatory response little affected by centre-surround inhibition or adaptation of the presumed ON and OFF channels. We include a stage for the recombination of ON and OFF channel signals, as yet untested by electrophysiological experiments, which enhances small target sensitivity. Finally, we show that this model leads to enhanced target discrimination, even when there are no relative motion cues between target and background.

2.2.3 Methods

2.2.3.1 Modeling

A model for an elementary small target motion detector (ESTMD) was created in Simulink (Mathworks), with image preparation and analysis tools programmed in Matlab (Mathworks). The term ‘elementary’ refers to a single unit that would be pooled to emulate the ‘position invariant’ nature of an STMD neuron (Nordström and O'Carroll, 2006). Each major component in the model (Figure 1C) is inspired by key stages in visual processing and will be discussed in detail later.

We do not attempt to emulate biophysical properties of cellular dynamics, e.g. compartmental modeling, nor are we developing a neural network representation. Rather we are building a numerical model based on linear and nonlinear spatial and temporal filtering and typical feed-forward signal processing methods. This approach allows for the model to be implemented in engineering applications.

The ESTMD model was tested using a series of panoramic images (Figure 2) (see Input Imagery) animated at a high temporal sample rate (5 kHz) to simulate continuous time. A 5x5 array of local ‘photoreceptor’ inputs was used to evaluate the response of the

central ESTMD (Figure 3). Because the input imagery is a circular panorama, continuous motion allows estimation of the output of this ESTMD for all horizontal locations on the image. The region of interaction was shifted vertically in 1° increments to build up a 2 dimensional representation of ESTMD outputs in a raster fashion (Figure 3). The stimulus was rotated at $90^\circ/\text{s}$ (within the optimal range for STMD neurons (Nordström and O'Carroll, 2006)) for two complete revolutions, with the first discarded, to avoid start-up transients.

2.2.3.2 Input imagery

To test for robustness of the model for discriminating targets embedded in visual clutter, a series of three panoramic images (Figure 2B-D), with a 72° vertical extent, were acquired from natural habitats (Brinkworth and O'Carroll, 2007). The 8-bit images were 2048×410 pixels. Although original panoramas were sampled as RGB, all simulations used the green channel only in order to approximate the spectral sensitivity of the fly photoreceptors that subserve motion processing (Srinivasan and Guy, 1990). A fourth image (Figure 2A) was obtained by combining the three natural images with ten others and averaging their phase and magnitude in the frequency domain (Brinkworth and O'Carroll, 2007). This combined image, whilst displaying a typical power spectrum, lacks hard edge-like contours found in many (but not all) natural images. This image acts as a control with respect to potential phase congruency components underlying motion detection mechanisms. The first stage of modeling emulates fly optics via spatial blurring (see Results: Photoreceptors), therefore reducing hard edges, including those of the targets. The Gaussian blurring is shown for two images (Figure 2E, 2F) with 20 scattered $1.6^\circ \times 1.6^\circ$ targets embedded. The targets effectively have varying contrasts, and are difficult to discern, revealing the challenging nature of this target discrimination task. All four of the images had power spectra showing an approximately $1/f$ dependence on spatial frequency f , which is typical of natural images (Tolhurst et al., 1992, Bex and Makous, 2002).

We created a second set of images, identical to the first, but into which black targets ($1.6^\circ \times 1.6^\circ$) were pseudo-randomly distributed, with each target centered on an ommatidial row. To improve computational efficiency, we inserted twenty such targets into each image. We maintain a 70° horizontal separation between the targets and a 6° vertical separation. This limits spatiotemporal interactions between the targets at any stage of the modeling, with a larger (in effect longer) horizontal separation required, as

this becomes the resultant temporal domain due to the panorama being rotated horizontally (influences are of longer-term adaptive components, e.g. photoreceptor dynamics). Because these targets become a feature of the image (i.e. there is no relative motion between targets and background) these simulations test the most demanding condition observed in physiological STMD experiments (Nordström et al., 2006).

Model simulation was run with a single control trial using the original panoramas without targets and 26 trials (for each image) in which different pseudo-random target distributions were used, allowing us to evaluate responses from a total of 26 x 20 target locations, across four images.

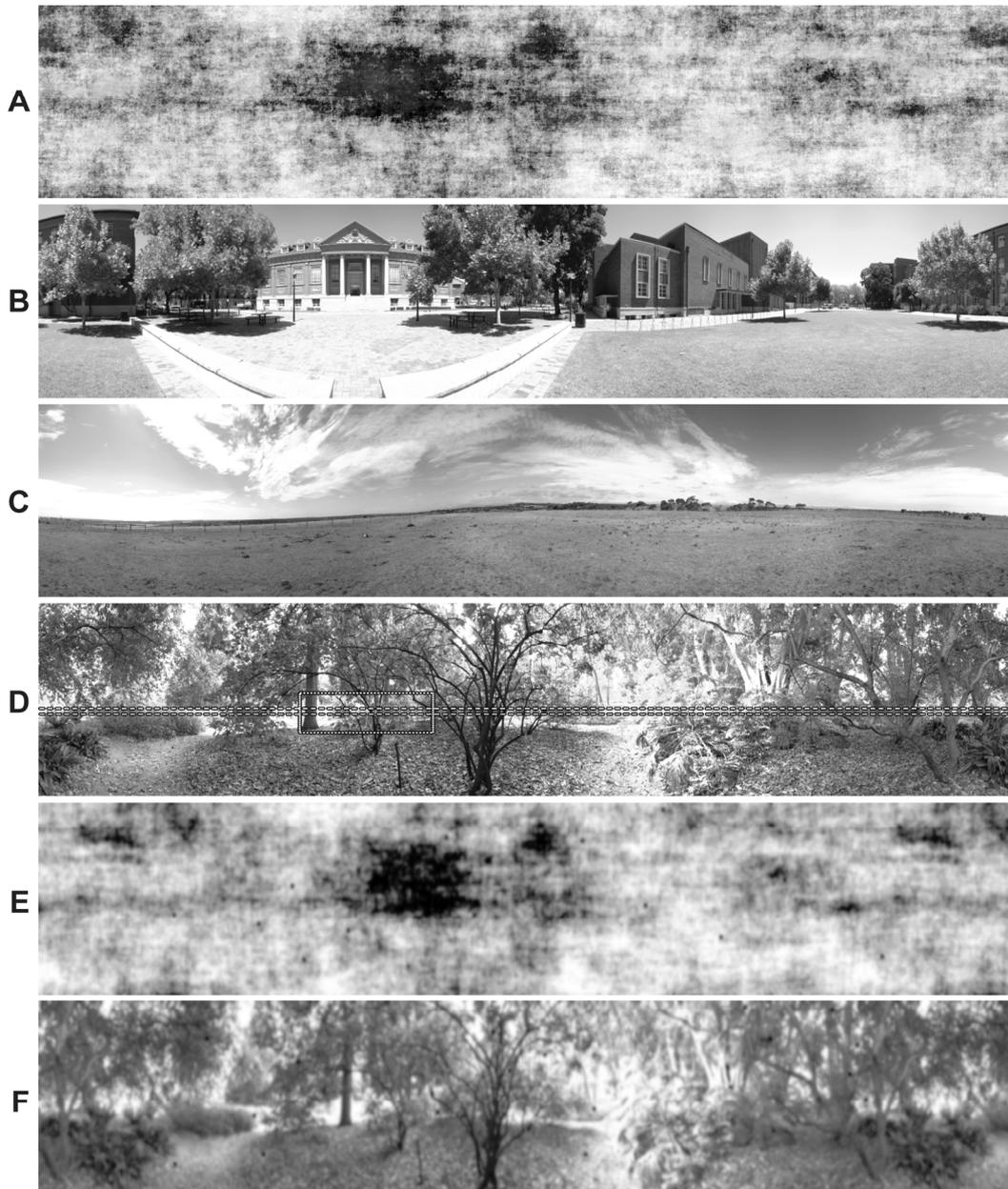


Figure 2-2. Input Images and Optical Blurring. Four panoramic images are used as model inputs to test target discrimination. The images display natural statistics with luminance intensity inversely proportional to spatial frequency (Tolhurst et al., 1992). Image (A) is composed of the average magnitude and phase of 13 natural images (Brinkworth and O'Carroll, 2007). Image (B) includes several man-made structural elements. Image (C) is relatively sparse, whilst (D) is a highly cluttered scene. The images are panoramic and extend 72° vertically. They have a resolution of 2048×410 pixels, with the 'green' channel of the RGB image (depicted here in grayscale) retained for further processing, approximating the spectral sensitivity of motion detection mechanisms in the fly visual system (Srinivasan and Guy, 1990). The row section highlighted in image (D) corresponds to the data traces of Figure 9. Images (E) and (F) are the optically blurred versions of images (A) and (D), including 20 pseudo-randomly scattered targets ($1.6^\circ \times 1.6^\circ$) in each image.

To analyze how effectively targets are discriminated a spatial image of the model output, at varying stages of processing, was reconstructed from the vertical columnar units, and binning of the horizontal time dimension (back into an equivalent 1° spatial domain). Target locations are determined taking the non-uniform lag into account. We determine ‘hits’ (above threshold output corresponding to target location) and ‘false positives’ (above threshold outputs not corresponding to target locations). This categorization is done for each image (Figure 2), at each processing stage, and across varying model output thresholds. By varying this threshold and plotting ‘hit’ rate (relative to total targets present in the scene) versus number of ‘false positives’, we constructed Receiver Operating Characteristic (ROC) curves.

In addition to the experiments using natural images, basic characteristics of the ESTMD were evaluated using a similar stimulus paradigm into which targets of varying contrast, height and velocity were animated against bright or mean luminance backgrounds.

2.2.3.3 Physiology

Flies (*Calliphora*) were either caught in the wild or reared in the laboratory under a natural day/night cycle. Insects were immobilized with wax. The back of the head was shaved, and a small hole in the cuticle was removed. Air sacs and other tissue were removed to provide clear access to the medulla. The brain was immersed with a Ringer solution: NaCl (130 mM), KCl (6 mM), MgCl₂ (4 mM), CaCl₂ (5 mM), with HEPES buffer at pH 7.0. Osmolarity was adjusted to 450 mM with the addition of sucrose. The fly was positioned to view a 200 Hz CRT monitor, mean luminance of 100 cd·m⁻². The visual stimuli were programmed in Python, using the VisionEgg stimulus software (www.visionegg.org).

Micropipettes were pulled from 1 mm (O.D.) thick walled alumina-silicate glass capillaries (SM100F-10, Harvard Apparatus Ltd.), on a Sutter Instruments P-97 puller, and filled with 2 M KCl or 2 M potassium acetate. Electrode resistances were typically 120-150 MΩ.

A wide-field, square-wave, flicker stimulus (1 Hz) induced opposing polarity potentials within the extracellular space. Intracellular recording from the RTC was identifiable by: a) a drop to resting membrane potential of approximately -60 mV; b) the full-wave rectification of the signal; c) depolarizing responses of 10-15 mV (graded), with ~10

mV spikelets. The data were sampled at 5 kHz during acquisition, using a National Instruments 16-bit ADC. Data analysis was performed offline with Matlab.

2.2.4 Results

We consider here in detail both the major stages of our model, and compare their outputs with known biological counterparts.

2.2.4.1 Photoreceptors

After target insertion we low-pass filter input images (Gaussian, half-width 1.4°) to mimic the spatial blur of fly optics (Figure 1C, LPF₁) (Stavenga, 2003). Luminance values sampled by “photoreceptors” at 1° spatial separation approximately match the resolution of *Eristalis* (Straw et al., 2006) and *Calliphora* (Land, 1997). For computational efficiency, we use rectangular sampling in a $5^\circ \times 5^\circ$ receptor patch, rather than emulating the hexagonal distribution of ommatidia (Figure 3). Photoreceptor transduction transforms the input luminance to membrane potential in a roughly logarithmic manner around an operating point determined by stimulus history (Laughlin, 1981a, Srinivasan et al., 1982, van Hateren and Snippe, 2001). Our model mimics this effect by transforming luminance values with a Lipetz function (Equation 1), with the exponent u set at 0.7, as in our earlier modeling of fly motion detection (Shoemaker et al., 2005).

$$\frac{x''}{x'' + x_0''} \quad (1)$$

To elaborate this Lipetz nonlinearity we include an ‘adaptation state’ with the mid-level parameter x_0 set as a first-order low-pass filtered version of x (time constant (τ) of 750 ms). Fly photoreceptor responses are temporally limited, with a corner frequency of 40-70 Hz (Laughlin and Weckstrom, 1993). To capture this, our modeling employs a static low-pass filter with corner frequency of 60 Hz (Figure 1C, LPF₂) following the Lipetz transform.

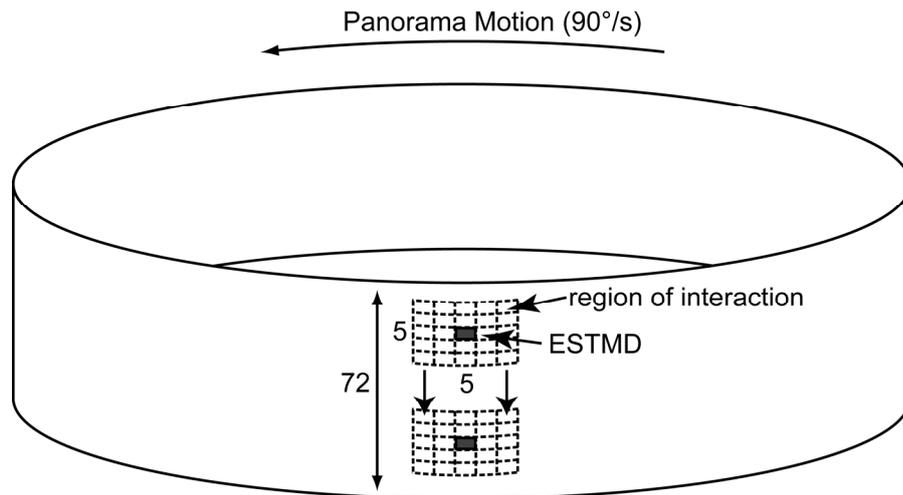


Figure 2-3. Panorama Rotation. A single model output has 25 ‘photoreceptors’ as inputs, each with 1° sampling separation (inter-ommatidial angle), thus represents a 5°x5° grid. The values representing luminance intensities at these locations vary over time as the panorama image is rotated past the ESTMD at 90°/s. Linear interpolation between pixels in the horizontal spatial domain results in higher temporal resolution (sampling at 5kHz). There is an ESTMD at each degree separation down the vertical column, therefore 72 in all, each with overlapping, feed-forward, receptive fields.

2.2.4.2 Large Monopolar Cells

While the role of LMCs in motion processing has been controversial, most research suggests that they are the ideal input to this pathway (Coombe et al., 1989, Douglass and Strausfeld, 2003, Keller, 2002). The LMCs have been shown to remove redundancy (Srinivasan et al., 1982) and maximize information transmission (van Hateren, 1992b) and they work as spatiotemporal contrast detectors, suitable for feature detection. Therefore, we implement an LMC-like spatiotemporal band-pass filtering on the photoreceptor output (Figure 1C). Spatial antagonism can be modeled as a recurrent inhibitory network (direct LMC to LMC inhibition), however, surround inhibition in a feed-forward manner, via a proposed surround ‘amacrine cell’ is equally plausible and is in accord with recent research on fly retina-lamina circuitry (Zheng et al., 2005). Our modeling comprises an amacrine cell that samples the surrounding nine photoreceptor outputs and subtractively inhibits the centre LMC (leaving a 30% DC spatial component). LMC spatial filtering dynamics are variable, dependent on overall light adaptation levels (Srinivasan et al., 1990b); however, our model parameters remain constant for computational efficiency. The inhibitory signal is delayed prior to the subtraction by application of a first-order low-pass filter (LPF₃, $\tau = 2$ ms), representing the time course of the amacrine cell signal spread (James, 1990). The LMC has band-

pass temporal characteristics, with low frequency roll-off below a few hertz and high frequency at ~80-100Hz, in light adapted conditions (Juusola et al., 1995). For our model, the LMC signal is temporally filtered (R-HPF₁) with a ‘relaxed’ first-order high-pass filter (one that passes a small DC component of 10%). This filter is characterized in the Laplace domain by the transfer function:

$$\frac{0.1*(10\tau s + 1)}{(\tau s + 1)} \quad (2)$$

where s is the Laplace variable and $\tau = 40$ ms. The LMC signal is inverted, to replicate the hyperpolarizing response to luminance increments observed in intracellular recordings (Järvilehto and Zettler, 1970).

2.2.4.3 Rectifying Transient Cell

Because electrophysiological data suggests that RTCs give little sustained response, unlike LMCs (Figure 4A), the signal is passed through an additional first-order high-pass filter (HPF₃, $\tau = 40$ ms). A half-wave rectification is performed to segregate ON and OFF channels of the input waveform, with the negative phase inverted in sign.

For each independent channel of the RTC, a signal representing an ‘adaptation state’ is formed by applying a nonlinear low-pass filter to the input signal with a fast onset, slow decay characteristic (Equation 3).

$$\begin{aligned} d/dt\{NLF\} &= (X - NLF)/\tau_1 \quad (X - NLF \geq 0) \\ &= (X - NLF)/\tau_2 \quad (X - NLF < 0), \end{aligned} \quad (3)$$

X designates the input, NLF the filter output, and τ_1 is set to 1 ms (LPF_{fast}) and τ_2 to 100 ms (LPF_{slow}). Such a filter is an approximation to plausible biophysical mechanisms, such as an interneuron with a long intrinsic membrane time constant and strong, ‘bursty’ inputs. This fast depolarizing, slow repolarizing signal is subtracted from the unaltered, pass-through version of the input signal.

In addition to this step, a second subtractive inhibition is applied based on the average of the surrounding input signals of the same channel polarity (surrounding ON subtractively inhibit the centre ON channel and similarly for the OFF channels). This is based on the surround inhibitory effect found in the ‘on-off’ cells (Jansonius and Van Hateren, 1993a). Unlike the previous parameters in our model, we do not have a

physiologically derived estimate for the strength of this inhibitory effect, and consider the scaling of the inhibitory signal a free parameter (INH) in our modeling and simulations. Alteration of this value can be used to tune the model to different size image features. We include a neural delay, modeled by a first-order low-pass filter (LPF₄, $\tau = 2$ ms), which is applied to the averaged and scaled surround inhibitory pathways.

The channels are then half-wave rectified to mimic a thresholded response (a nonlinearity seen in many spiking neurons). The resultant channel signals are passed through a ‘neural delay’ smoothing filter ($\tau = 2$ ms). This smoothing better represents the temporal response dynamics seen in the physiological RTC.

The final stage of processing is a recombination of the ON and OFF channels to form a single output corresponding to the ESTMD response. The simplest operation to achieve this would be a straightforward sum of the two output signals. However, we consider an operation that enhances selectivity for small, dark targets. A delay operator $D[*]$, consisting of a low-pass filter (LPF₅), is applied to the OFF channel prior to recombination with the undelayed ON channel. For generality, we took a phenomenological approach to this recombination allowing second-order as well as linear interactions:

$$Output = a \cdot ON + b \cdot D[OFF] + c \cdot ON \cdot D[OFF] \quad (4)$$

In our simulations, we consider primarily the purely linear case ($c=0$), which we refer to as ‘RTC’, and the second-order case ($a=b=0$), referred to as ‘ESTMD’. Note the formal similarity of the second-order structure to the correlational or Hassenstein-Reichardt elementary motion detector (Hassenstein and Reichardt, 1956). However, in this case the correlation operates on rectified signals of opposite polarity from the same spatial location, rather than signals from spatially neighboring locations. In this form, although tuned primarily to small contrasting features, this rectification of polarities resembles models proposed to explain selectivity for expanding edges in ‘looming’ motion detectors such as in the locust LGMD/DCMD (Rind and Bramwell, 1996, Rind and Simmons, 1992).

Although STMDs respond better to black targets (Nordström et al., 2006) and light target sensitivity is not modeled here, a symmetric correlation operation could be

established for a white target detector by interchange of the signal roles in Equation 4. This would provide white target sensitivity by correlating a delayed ON channel with an undelayed OFF channel. A detection mechanism for targets of both contrast polarities (light and dark) would involve summing these two versions or having any weighted combination of the above terms (both linear and second-order).

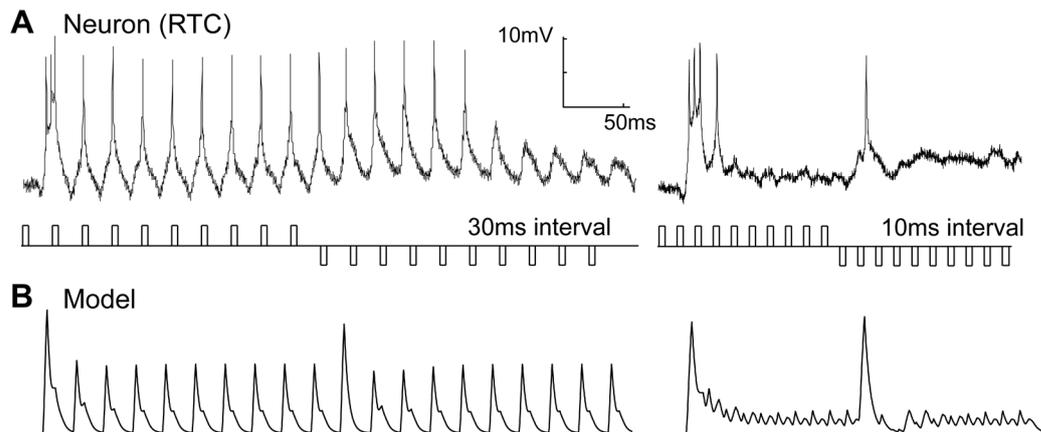


Figure 2-4. Rectifying Transient Cells (Independent Adaptation). (A) Intracellular recording from a rectifying transient cell shows independent adaptation to ‘on’ and ‘off’ contrast changes. Pulses of equal, but opposite, contrast polarity are of 5 ms duration and separated by varying intervals. At a pulse interval of 30 ms (left trace), little adaptation can be seen (but the response to the first pulse is strongest). At 10 ms interval, temporal adaptation limits the response *until the polarity of the contrast reverses*, which then produces an unadapted response. A single electrophysiological recording is shown for clarity (from N=5). The model version (B) output is similar, capturing the important functional characteristics of the RTC.

2.2.4.4 Comparison of model responses to fly RTCs

We compare our recordings of the RTC in the medulla of the blowfly to our modeled responses. The intracellular recordings (Figure 4A) show independent adaptation to contrast increments and decrements, as seen in ‘on-off’ type cells (Jansonius and van Hateren, 1991, Osorio, 1987). Figure 4A shows an experiment with a train of contrast pulses at two different frequencies. At 30 ms separation, the neuron recovers to produce graded depolarization in response to each pulse. When the separation is reduced to 10 ms, the adaptation suppresses the response to the stimulus. However, when the contrast polarity is reversed (from contrast increments to decrements), an unadapted response is observed before the neuron again rapidly adapts to the new polarity stimulus.

2.2.4.5 Temporal responsiveness

Although our model captures the basic behavior of the biological RTC, the incorporation of an LMC-like input stage is somewhat contradicted by earlier work suggesting the frequency response (to sinusoidal stimuli) of ‘on-off’ units rolls off above 12 Hz (Jansonius and van Hateren, 1991), while the LMCs have a much higher corner frequency (Juusola et al., 1995). Jansonius and van Hateren (Jansonius and van Hateren, 1991) suggested this apparent low-pass characteristic is simply a result of the rapid adaptation that occurs at higher stimulus frequencies (as seen in Figure 4A); it is possible that the unadapted system has a much higher temporal acuity than this result would suggest.

To test this hypothesis we used a ‘doublet’ stimulus consisting of a pair of pulses (‘on’ followed by ‘off’). Whilst not strictly containing energy at a single frequency, this stimulus allowed us to construct transfer functions for the RTC to a single stimulus cycle, thus avoiding the influence of adaptation. The response power is calculated as the mean-square value until the neuron returns to within 5% of the resting membrane potential. As can be seen in our physiological data (Figure 5A, dashed line, squares), the response of the medulla RTC to the doublet stimulus has a peak at high frequency (~50 Hz). The RTC still responds with 85% maximum at 100 Hz, the highest frequency doublet that we could generate on our 200 Hz stimulus display. The model RTC (Figure 5A, dashed line) gives a similar temporal responsiveness. The RTC frequency response is a good match for that obtained by Fourier transforming the linear kernel for fly LMCs using white noise stimuli (Figure 5A, solid line) (1990). Interestingly, if we simulate the earlier experiments of Jansonius and van Hateren (Jansonius and van Hateren, 1991) with a wide-field sinusoidal stimulus (Figure 5B, dashed line), we obtain a curve that rolls off at a much lower frequency, consistent with their experimental data (reproduced in Figure 5B, solid line). Our model rolls off more sharply at low frequency, likely due to the pure nature of our high-pass filter (HPF₃) and because the non-linearity introduced into their extracellular recordings by the thresholding mechanism for spike generation may lead to overestimation of weak responses. We conclude that the apparent low-pass nature of the ‘on-off’ cell frequency response was, as hypothesized, a result of adaptation (Jansonius and van Hateren, 1991), and that the response to transient as opposed to stationary stimuli reflects a much more rapid temporal response capability. Also, the inclusion of an LMC-like input stage

in our model is supported by the very similar temporal characteristics of the LMC to the fly RTC.

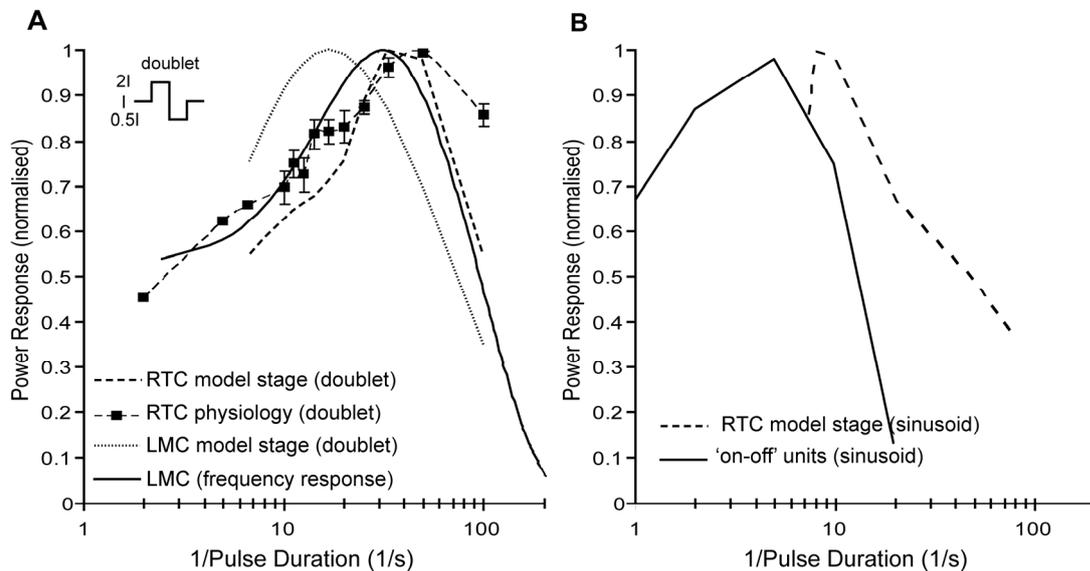


Figure 2-5.. Temporal Responsiveness. (A) The response of the physiological RTC to a stimulus doublet (2, 5, 6.6 Hz N=2, no error bars; others N=6 (flies) mean \pm SEM). This RTC transfer function peaks at \sim 50 Hz (dashed line, squares) and is still responsive at the highest stimulus frequencies, which are limited by the CRT refresh rate. We simulate doublet input and show that the model RTC frequency response is comparable to the physiological correlate (dashed line). The response to the doublet at the LMC stage of the model is also shown (dotted line). Frequency response properties of fly LMCs obtained via white noise analysis (James, 1990) is also plotted for comparison (solid line). It should be noted that the RTC and LMC response characteristics show a similar temporal responsiveness.

(B) Previous analysis of ‘on-off’ units in the fly lamina (Jansonius and van Hateren, 1991) showed poor temporal responsiveness (peak at \sim 6 Hz) (reproduced here, solid line) and our model shows a similar shift in response to the non-optimal sinusoidal stimulus (dashed line).

2.2.4.6 Contrast Sensitivity Function

The high-pass nature of the RTC data (and as captured by our model) we expect to form an ideal basis for a neural pathway for small target detection as the signal from the passing target boundaries provides a near optimal transient stimulus, with no spatial antagonistic suppression that would occur with larger features.

We determine the model response to a small target ($0.8^\circ \times 0.8^\circ$) as a function of target contrast and compare it with that induced by wide-field flicker stimuli (Figure 6). As the target is below the size of a single ommatidium, an effective neural contrast is calculated by the convolution of the target with the optical blurring filter (half-width

1.4°) (Nordström et al., 2006). Even very low contrast discrete targets induce a model response over 10 times higher than that of the wide-field flicker stimulus (compare at equivalent contrast Figure 6A with 6B, dashed lines). We also plot reproduced STMD responses to targets of varying contrast (Figure 6A, squares). However, it is important to note that these responses were to 0.8°x0.8° targets (50°/s) moving on complex moving backgrounds (45°/s) (Nordström et al., 2006).

Physiological data for the low contrast sensitivity of ‘on-off’ cell responses to wide-field flicker (Jansonius and Van Hateren, 1993a) is well explained by the model (Figure 6B). The divergence seen between the model and neuron recordings at higher negative contrast is expected, since we make no attempt to account for saturating nonlinearities in neural components that would be expected in the biological system. Interestingly, the RTC model stage also produces a reasonable explanation for the near threshold contrast sensitivity of higher order STMD neurons (Figure 6A, squares) (Nordström et al., 2006).

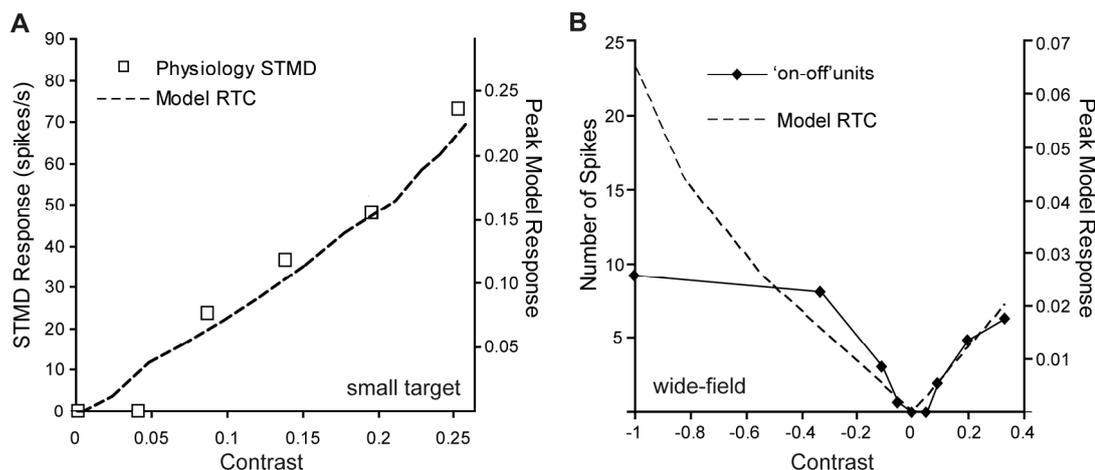


Figure 2-6. Contrast Sensitivity Function (Small Target and Wide-field). The contrast sensitivity function is calculated from the peak model responses to varying contrasts of small targets (0.8°x0.8°) moving at 50°/s on a mean background (RGB 0.5) (dashed line). Also plotted (squares) are physiological STMD responses to a 0.8°x0.8° target, moving at 50°/s. However, this stimuli included a complex moving background (mean luminance 150 cd·m⁻², 45°/s) (Nordström et al., 2006). For the model we measure contrast as the effective neural contrast. For the physiological data the contrast values represent average Michelson contrasts as the targets transverse a complex moving background (Nordström et al., 2006). (B) Reproduced responses of ‘on-off’ units to wide-field contrast steps of 500 ms duration (solid line) (1991). For comparison, we plot the model RTC responses to a simulation of this wide-field visual input (dashed line). We note that low contrast sensitivity is observed in the model output due to spatial antagonistic interactions and this could be a plausible explanation for low contrast sensitivity in the ‘on-off’ units. The model responses in (B) are less than 1/10 those seen in response to small targets of equivalent contrast (A).

2.2.4.7 Target height tuning

A feature of our ESTMD model is the inclusion of second-order spatial (lateral) inhibition by neighboring RTCs and a temporal cross correlation of the outputs of local ON and OFF pathways which form a ‘matched filter’ for both the spatial and temporal characteristics of small, moving features.

By analogy to models for direction-selective motion detectors where wide-field optic flow can be deduced by summing output of local elementary motion detectors, we use the term ‘elementary small target motion detector’ (ESTMD) for this stage. Responses of higher-order STMDs should be easily explained by simply summing across a weighted array of such ESTMDs to produce receptive fields of varying size (as observed in electrophysiological recordings from the lobula) (Barnett et al., 2007, Nordström et al., 2006) whilst retaining position invariant selectivity for small features (Nordström and O’Carroll, 2006). To confirm whether our model displays size selectivity, we estimate responses to discrete moving targets of different length (i.e. extended orthogonal to the direction of motion). Figure 7 shows that the ESTMD stage of our model provides an excellent fit to the data published for lobula STMD neurons (Barnett et al., 2007). Note that while LMCs act to maximize information to the higher order pathways by enhancing edge-like features, the very sharp suppression of responses to targets above a few degrees in size that characterizes both model and neuron responses cannot be explained by the simpler spatial antagonism of LMCs (Figure 7, dashed line).

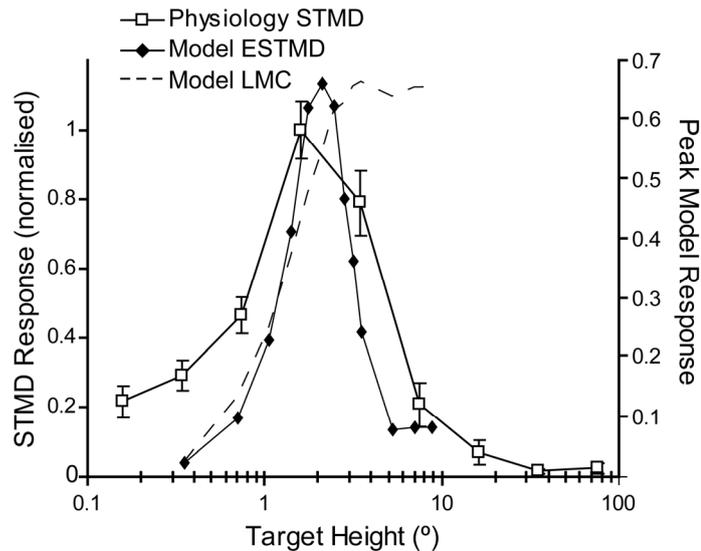


Figure 2-7. Target Height Tuning. The curve shows model ESTMD responses to targets of varying height (0.8° wide black targets on white background, moving at $50^\circ/\text{s}$). Physiological data from STMD neurons in the hoverfly for the equivalent target parameters and background is reproduced (Barnett et al., 2007). The model is selective for targets of less than a few degrees height with the suppression to the right of the peak determined by the strength of the lateral inhibition between channels. The model response to targets at the LMC stage (no units) shows that the responses remain at maximum as the target is extended vertically (max height shown of 10°), i.e. the LMC is not target selective. This highlights that a second-order spatial antagonism is required for target selectivity.

2.2.4.8 Velocity tuning

An important aspect of the second-order configuration of our model is its inherent similarity to a Reichardt correlator (Hassenstein and Reichardt, 1956) such that the velocity dependence in response to small moving targets is essentially the same. The responses to a $0.8^\circ \times 0.8^\circ$ moving target (Figure 8) represents a typical velocity tuning curve as obtained from a delay-and-correlate-type model. The position of the peak response is dependent on the time constant of the delay filter $D[*]$ (LPF_5). For comparison, we plot the velocity tuning curve seen in STMD neurons (Nordström and O'Carroll, 2006). We have not attempted to specifically fit this data (nor in the target height tuning) and note that differences in the broadness of the tuning curves could reflect additional compressive nonlinearities which we have not attempted to account for in this model.

Although the ESTMD model provides a good account for the basic tuning properties of STMDs, it is not unique in this respect. Other STMD models (Egelhaaf, 1985c, Higgins and Pant, 2004) should also be able to explain both contrast sensitivity and velocity

tuning. However, our key finding is that the unique adaptive component of the RTC inputs to our STMD can also explain the otherwise enigmatic finding that STMDs can respond to features embedded in clutter, but without *relative motion* cues (Nordström et al., 2006).

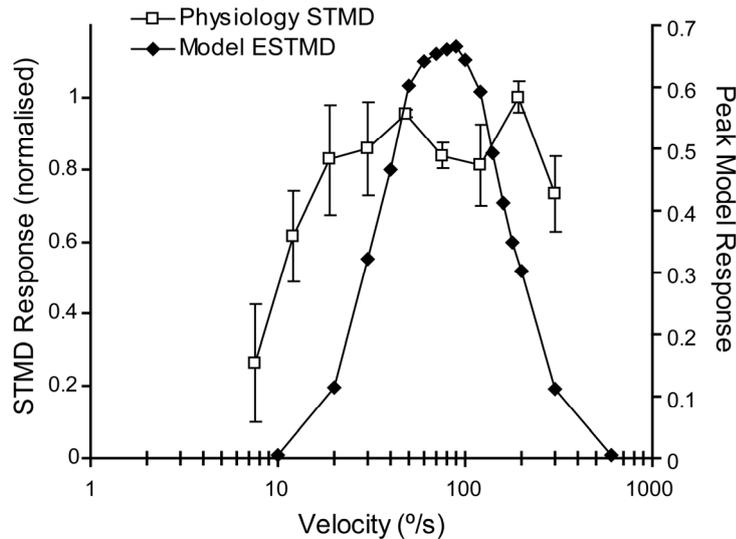


Figure 2-8. Velocity Tuning. Model ESTMD responses to black moving targets ($0.8^\circ \times 0.8^\circ$) on a white background at varied velocities is shown in comparison to the physiological data of STMD neurons to the same visual stimulus from hoverflies (Nordström and O'Carroll, 2006). The model output exhibits typical velocity tuning as observed in correlation-type motion detection mechanisms. The tuning of the model parameters (particularly, the OFF delay filter time constant) determines at which point the velocity response peaks. The broadness of the tuning curve may be extended and shaped via the addition of a final saturating nonlinearity, not included in this model.

2.2.4.9 Responses to targets in clutter

Figure 9 shows a single output row at each stage of the model, in response to a panoramic image in which a small target is inserted (the image row is delineated in Figure 2D). We selected this row to illustrate the effect of the key stages of the model in enhancing target responses, whilst rejecting other high contrast features.

Photoreceptor dynamics encode a large luminance range into the limited dynamic range of the neuron (Brinkworth et al., 2006, van Hateren and Snippe, 2001). Our inputs have already emulated a similar process via a form of global gain control inherent in digital camera processing (Figure 9A-B). The LMC output (Figure 9C) with spatiotemporal high-pass filtering, enhances contrast boundaries in both space and time. The OFF fast temporal adaptation (Figure 9D, solid line) suppresses textures and signals larger ‘breakthrough’ contrast changes.

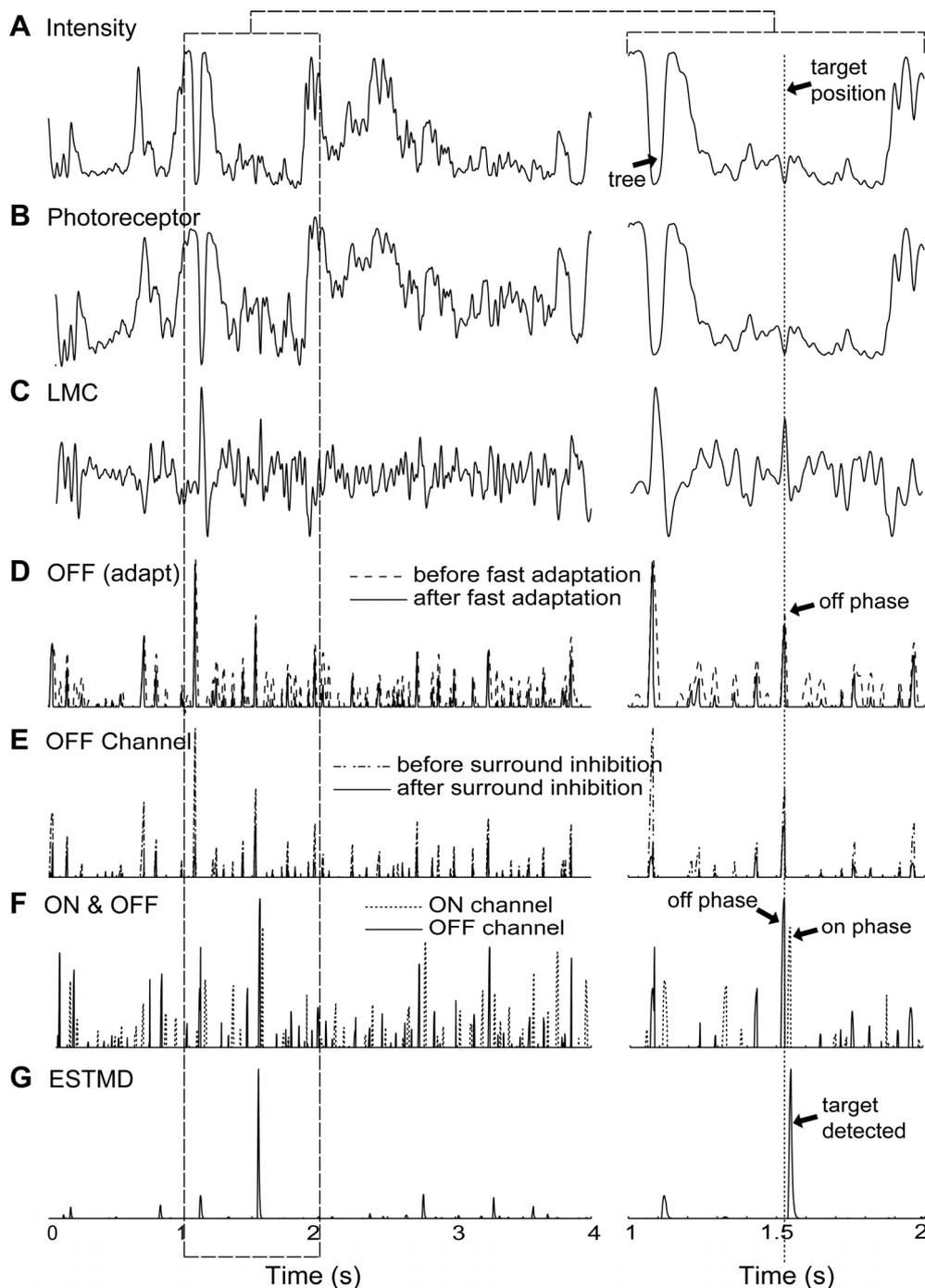


Figure 2-9. Sample Data Traces. The traces show the mode outputs at various stages of processing over the second complete revolution of the scene (at $90^\circ/\text{s}$). The right hand trace shows a magnified version of the period between 1 and 2 seconds. The y-axes are unit-less model outputs. The input intensity (A) is delineated in Figure 2D. Photoreceptor dynamics (B) encode a large luminance range into the limited dynamic range of the neuron. The LMC (C) exhibits spatiotemporal high-pass filtering, enhancing contrast boundaries. The temporal adaptive mechanisms within the independent channels suppress rapid texture variations but signal a novel contrast change (D). Surround antagonistic interactions limit the spatial extent of this type of signaling (E) and a final linear or quadratic recombination of the channels (F) signals the presence of a dark moving target (G).

The surround inhibition ensures that this effect is spatially localized. Note that the response to the tree trunk ($t = 1.1$ s), which also has a novel ‘off’ shortly followed by an ‘on’ contrast boundary is suppressed as a consequence of second-order spatial inhibition (Figure 9E, solid line). The ON channels (not shown) show similar characteristics. Finally, the OFF channel (Figure 9F, solid line) is temporally delayed and correlated with the undelayed ON channel (Figure 9F, dashed line) to signal target-like events (Figure 9G).

Figure 10 shows ROCs for the four panoramic images, at a velocity of $90^\circ/\text{s}$. During the pseudo-random distribution, some targets are scattered onto backgrounds of the same luminance (as the target) such that they lose all defining characteristics. In image D (Figure 2), the most highly textured scene, it is difficult for the human observer to detect the scattered targets. Image C is extremely sparse and LMC filtering is enough for successful target discrimination (Figure 10C). Across the varied scenes, both linear (RTC) and quadratic (STMD) processing have improved the discrimination of targets as revealed by the shift to the upper left corner of the ROC curve (Figure 10 A-D). The limited number of false positives in the final model output suggests that target-like structures are rare in these natural image scenes.

These results show that a highly nonlinear filter (derived from the plausible biological components) exploits the spatiotemporal statistics of the moving target within its immediate surround. The statistics required are as follows 1) a small duration of time (~ 50 ms) in which contrast changes do not exceed that of the upcoming target, therefore providing an unadapted ‘off’ phase. This provides ‘distinctiveness’ to the start of the dark feature. 2) An unadapted ‘on’ phase, which is inherent in the non-changing texture of the dark target. 3) These same characteristics, i.e. unadapted, opposite polarity, contrast changes, *to not be present* in the immediate surround. If this third characteristic were relaxed, the detector would be sensitive to a similar width/velocity profile as the target, though not suppressed by the height of the feature, i.e. the detector would also be stimulated by a vertical ‘bar’ stimulus.

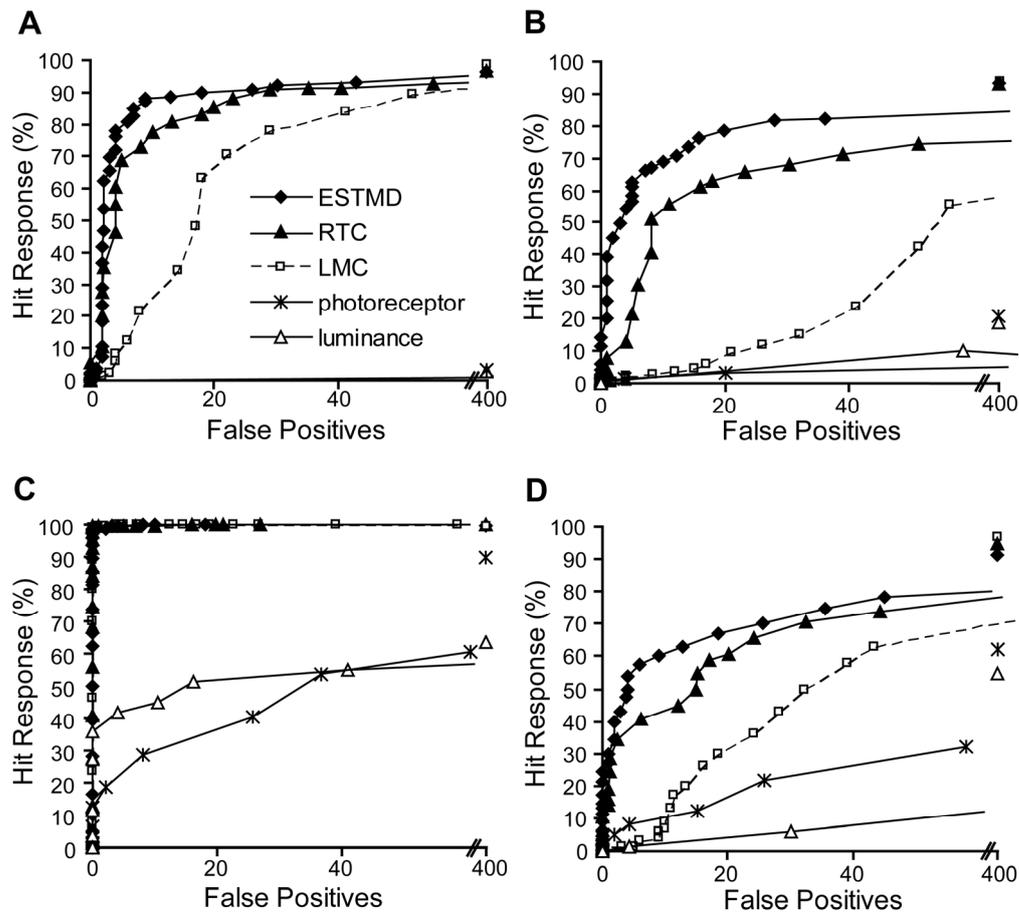


Figure 2-10.. Natural Image Panoramas (ROC curves). Twenty targets ($1.6^{\circ} \times 1.6^{\circ}$) pseudo-randomly scattered (26 trials) on four panoramic images with the simulation run at a velocity of $90^{\circ}/s$. (A) Receiver Operating Characteristic (ROC) curves are shown for a scene with averaged natural statistics. The addition of RTC-type processing (solid line, triangle) to the LMC (dashed line, square) shifts the ROC curve to the upper left, revealing enhanced target discrimination. The quadratic (ESTMD) version of the model (solid line, diamond) shows further improvement via the multiplicative interaction of the delayed OFF and undelayed ON channels. (B) The LMC stage has a large number of false positives, due to high contrast, man-made features (which the RTC-type processing can discriminate). (C) This image is sparse so the targets can be readily discriminated by the LMC processing alone. (D) A highly textured scene, with many scattered targets losing defining characteristics, however, the target discrimination is still improved. Error bars are within symbol representation, therefore removed for clarity.

2.2.4.10 Relative motion

Intuitively, the ESTMD model is responsive to the motion of the contrast boundaries across the detector inputs. Relative motion between target and background will have an effect on ESTMD responses, as it alters the temporal statistics (dependent on background velocity) that establish the adaptation states of the independent channels. We tested this by varying the background motion with a constant target velocity of $90^\circ/\text{s}$ (Figure 11).

Depending on background speed, we varied initial background position so that we could analyze target response at the same spatial juxtaposition of target and background (target size of $1.6^\circ \times 1.6^\circ$). Hence, data for a background speed of $+90^\circ/\text{s}$ effectively represents the scenario in the other ‘no relative motion’ panorama simulations. We repeated this test in 100 distributed locations across the four panoramic images. We show that the ESTMD target responses are robust across the tested range of relative motions and the results confirm that the response improves when there is some relative motion, reaching a peak when the background speed is close to zero.

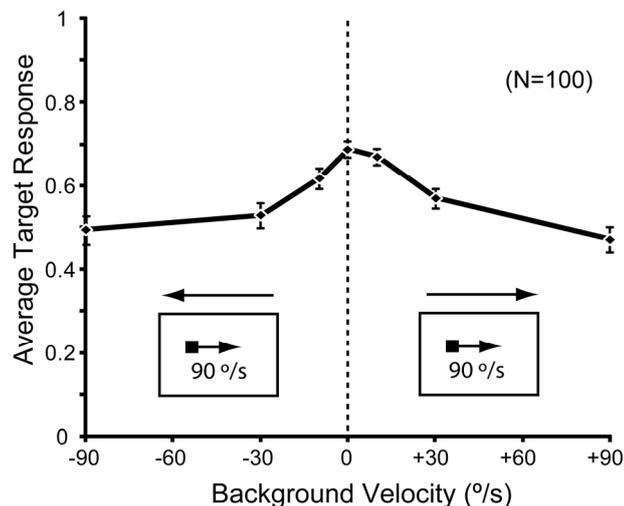


Figure 2-11. Relative Motion. Results of simulations carried out for 25 targets in each of the 4 images, where background speed and direction was varied. Target speed was constant at $90^\circ/\text{s}$ (left to right, as indicated by pictograms). Data are mean \pm SEM, $N=100$.

2.2.5 Discussion

By recording from a cell in the fly medulla (the RTC) we have been able to determine their quick temporal responsiveness to transient stimuli. This has provided parameters, such as adaptation time constants, that form the basis of our target detection model. We have compared responses of the model, e.g. contrast sensitivity, height tuning, and velocity tuning, to those observed in STMD physiological recordings and find that they are well matched.

We have shown that linear systems analysis of steady-state responses is not an appropriate method for characterization of the relevant neural responses (Figure 5). In this case, when we consider the quickly adapting transient component of the RTC signal, we find the temporal responsiveness is well matched to presumed early components of the visual system (the LMC). We also observe that apparent contrast insensitivity of a neural system may be the result of wide-field antagonistic interaction. In our modeling, due to a neural delay in the surround interactions, the naturally ‘unrealistic’ wide-field flicker can transiently pass through the system, however with low contrast sensitivity the result. In comparison, the response to the limited spatial extent of a small target is very contrast sensitive.

Larger wide-field STMD neurons display a position invariant receptive field that typically spans many ommatidia (Nordström and O'Carroll, 2006). This presumably requires a pooling of the outputs of many subunits from earlier stages of processing. In fact, because some STMD neurons are weakly direction-selective (Barnett et al., 2007, Nordström et al., 2006, Nordström and O'Carroll, 2006), rather than a summative pooling of ‘non-directional’ subunits, some type of higher-order spatial facilitation may take place. Weak direction selectivity could be built into our modeling via asymmetry in the inhibitory surrounds or via a higher order spatial facilitation during this pooling stage.

Unfortunately, RTC neurons are small and intracellular recording times are limited in duration. We have, to date, not been able to establish the morphology of the neuron via dye-filling techniques nor can we examine more time intensive spatial characterization stimuli. Nevertheless, we are confident from our dissection technique and precise control of the location of the pipette that our regular recordings from the RTC are from the medulla. However, we cannot be certain if they are intrinsic to the medulla, or if

they reside elsewhere and project to, from, or via the medulla. There is the possibility that they may be the termination of the fibers identified by Arnett (Arnett, 1971, Arnett, 1972) or a later postsynaptic element that has inherited the properties as seen in the projections from the lamina. Although our biological investigation of the RTC is limited, the aspects of computation that form the basis for our small target modeling has been well established in the work of Jansonius and van Hateren (Jansonius and van Hateren, 1991), Osorio (Osorio, 1987) and now again in this present research.

2.2.6 Conclusion

Our approach to modeling has provided a solution to the initially perplexing issue of how the STMD neuron responds robustly to target motion, even when there is no relative motion cue of the target to the background (Nordström et al., 2006). We have seen that this problem is solved by incorporating properties of the rectifying transient cell in the target detection pathway. This is an attractive solution, as our highly nonlinear matched filter is computationally less intensive than complex segregation of transparent motion fields, required for relative motion calculations.

2.2.7 Acknowledgements

Panoramic Images courtesy of Dr Russell Brinkworth. We thank the manager of the Botanic Gardens of Adelaide for allowing us to collect insects. We also thank past and present members of the O'Carroll lab for contributing discussions, especially Dr Russell Brinkworth, Dr Karin Nordström and Paul Barnett.

CHAPTER 3.

ENGINEERED TARGET DISCRIMINATION DERIVED FROM INSECT PHYSIOLOGY

3.1 CONTEXT

This chapter was published in IEEE, Intelligent Sensors, Sensor Networks and Information Processing (ISSNIP).

There is a large degree of overlap between the publication in this chapter, and the one preceding (see 2.2). However, the ISSNIP publication was formatted and written for an engineering audience, and this adopts a different approach to the article published in PLoS ONE. The later was written with much more extensive background and supporting physiological data for a broad readership.

This degree of overlap is not the case in the chapters following this one.

3.2 BIOLOGICALLY INSPIRED SMALL TARGET DETECTION MECHANISMS

S. D. WIEDERMAN¹, P.A. SHOEMAKER², D.C. O'CARROLL¹

¹*Discipline of Physiology, School of Molecular and Biomedical Science, The University of Adelaide, Adelaide, Australia*

²*Tanner Research Inc., Monrovia, California, USA*

PUBLISHED: Wiederman, S., Shoemaker, P. and O'Carroll, D.C. (2007) Biologically inspired small target detection mechanisms. Proceedings IEEE, Intelligent Sensors, Sensor Networks and Information Processing (ISSNIP) Conference, pp 269 - 275

3.2.1 Abstract

A challenging engineering problem is the ability to detect and discriminate small moving objects against complex moving backgrounds. An evolutionary priority has tuned this capability within fly vision systems, so that they may detect, track and chase other flies, either for territorial or mating purposes. This is achieved within the confines of a light weight, low power system of less than a million neurons. Using in-vivo, intracellular, electrophysiological techniques, our lab has recorded from neurons within the fly brain that respond to a visual stimulus of small moving targets, even when included within a complex surround. Intracellular responses are still present when the velocities of the target and background are matched, ruling out relative motion cues as the definitive factor in this target discrimination task.

Intracellular recordings at an earlier stage of the visual pathway show neurons with rectified, transient responses that have independent adaptation to stimulation with contrast increments versus contrast decrements. These Rectifying Transient Cells (RTC) have an 'on' and an 'off' channel that we hypothesise would serve well as an intermediate stage in the detection of small moving targets.

To test the feasibility of this hypothesis, we have developed a model that includes the filtering characteristics of early visual processing, followed by the rectification and

independent adaptation properties of the RTC and finally a temporal nonlinear facilitation between a delayed ‘off’ channel, and an undelayed ‘on’ channel. The model output shows an enhancement in target discrimination, without the need for relative motion cues, as seen in physiological experimentation. Characteristic properties of neuronal responses from small target detecting cells are well matched by the outputs of this model.

3.2.2 Introduction

Flies have the capability to detect and track small moving objects, such as other flies, often whilst both travel at high velocities and sometimes against complex moving backgrounds (Collett and Land, 1975, Wehrhahn, 1979). This is a challenging task, considering the low spatial acuity of their visual system which has an angular resolution of approximately 1° (Stavenga, 2003). This capability has led researchers to investigate the neural elements behind the processing of such a difficult visual task. Neurons are studied, via intracellular recordings, at various stages of the visual pathway, in an attempt to understand and model the components of this processing. These electrophysiological techniques have examined cells that are most sensitive to small moving objects in a range of insects: blowfly (Wachenfeld, 1994), dragonfly (O'Carroll, 1993), fleshfly (Gilbert and Strausfeld, 1991) and hoverfly (Collett and Land, 1975).

We recently described neurons within the optic lobes of the hoverfly *Eristalis tenax* that are exquisitely selective for small targets of less than a few degrees (Nordström et al., 2006). These neurons are referred to as Small Target Motion Detectors (STMD) and respond in an invariant manner to target position within an area of the visual field (receptive field) and are often direction selective.

A characteristic of the STMD neuron is that their responses maintain small target selectivity even with the presence of a complex moving background. Most intriguingly when the velocities of the target and background are matched responses persist.

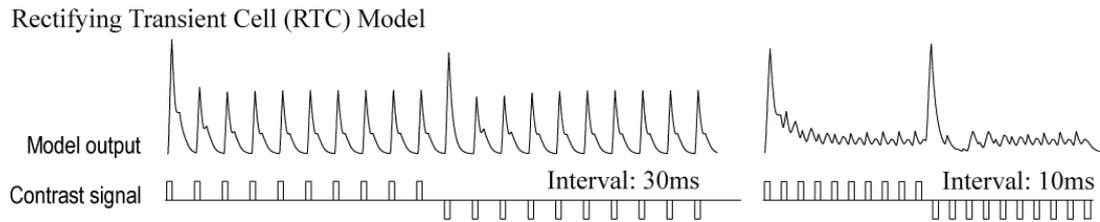


Figure 3-1. Responses of Model Rectifying Transient Cell. The model responses include characteristics of the RTC neuronal response. The output is transient and rectifies the signal (same response to contrast increments and decrements). Within short interval durations the fast adaptive mechanism reduces model output. This adaptation is independent for contrast increments versus contrast decrements as seen in the unadapted ‘off’ response, at the first contrast decrement pulse, in the 10ms trace.

This result suggests that the spatial statistics of the target with respect to the background is an effective discriminant, irrespective of any additional role for relative motion cues.

The identification of neural elements that may lie along the visual pathway of the STMD, and are possible to record from intracellularly, aids in our attempts to understand the neural computations underlying processing of the STMD. We have experimented with, and modelled, a neuron that we hypothesise may play such a role.

Previous experimentation has identified neurons within a variety of invertebrates that are characterised by their rectifying and transient nature (Arnett, 1971, Jansonius and van Hateren, 1991, Osorio, 1987). In the fly visual system, these cells, referred to as ‘on-off units’, revealed independent adaptation to contrast increments versus contrast decrements. There is also supporting evidence for the presence of centre-surround antagonism between such units (Jansonius and Van Hateren, 1993a).

We have recorded intracellularly from neurons, that we refer to as Rectifying Transient Cells (RTC), within a middle stage of the fly visual pathway (the medulla), which also show similar response characteristics to the on-off units.

To examine the plausibility of our hypothesis that the RTC is well suited to the task of small target motion detection, we have modelled the STMD pathway, based on early vision and RTC processing. The model outputs are then examined for enhancement of target discrimination when provided with inputs of moving natural scenes embedded with small targets.

3.2.3 Modelling & Simulation

3.2.3.1 Overview

Image preparation, model simulation and analysis tools were written and implemented in Simulink/Matlab (Mathworks, Natick, USA). The model was designed using parameters consistent with those found in the biological system. An overview of the model is seen in Fig 2. We refer to the model output as an eSTMD, standing for an ‘elementary STMD’ as the physiological correlate STMD would be a number of these eSTMD subunits combined together to provide a broad, position invariant receptive field.

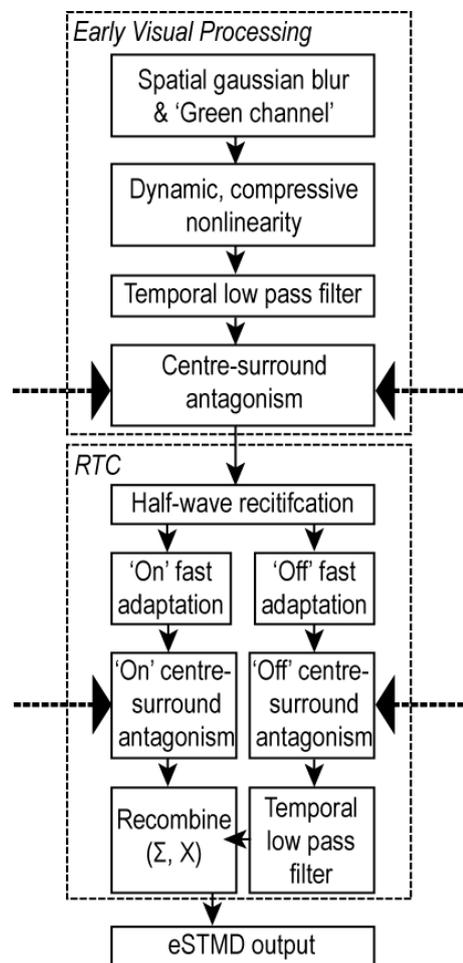


Figure 3-2. Overview of the ESTMD model. Early visual processing (optics, photoreceptors, LMC) and the RTC representation. Filter parameters are chosen so as to match the dynamics we observe in electrophysiological experimentation of the fly brain. Further details are found in text.

3.2.3.2 *Input Images*

A series of four natural panoramic images with a 360° horizontal and 72° vertical extent served as inputs to the model. These images could be animated in simulated yaw rotation past a column of STMD processors at varying velocities. To improve computational efficiency, twenty individual black targets (1.6°) square were pseudo-randomly distributed on the images, a minimum separation ensuring that there was no possibility of spatiotemporal interaction between the targets at any stage of model processing. Each simulation was repeated 26 times with new target locations (520 targets total for each image). During the simulation of panorama motion (yaw) the targets are *fixed* to the background, therefore there were no relative motion cues. The results were obtained using a panorama, and therefore target, velocity of 90°/s. Only the green channel of the RGB image was retained, so as to approximate the green spectral sensitivity of motion processing mechanisms found within the fly visual system.

3.2.3.3 *Early Visual Processing*

Input images were spatially blurred with a Gaussian low pass filter (half-width 1.4°) to emulate the spatial blur seen within the optics of the fly visual system (Stavenga, 2003). Luminance values from the input images were spatially sampled at 1° to represent the distribution of facets in the eye of the fly (Straw et al., 2006). An approximation in our model is that we used a rectangular distribution of sampling points, rather than the hexagonal distribution seen in the fly compound eye.

The photoreceptor transduction mechanism transforms luminance inputs in a logarithmic manner around an operating point determined by the past stimulus (van Hateren, 1997). This was modelled with a Lipetz function relating membrane potential V_m to light intensity x , as seen in equation 1.

$$V_m = \frac{x^u}{x^u + x_0^u} \quad (1)$$

The exponent u was set to 0.7 (Shoemaker et al., 2005). The mid-level parameter x_0 was set as a moving average of previous stimulus values, using a first order low pass filter ($\tau=750\text{ms}$). The temporal characteristics of the photoreceptor were modelled with a first order low pass filter ($\tau=2.5\text{ms}$).

The next stage of processing is represented by the Large Monopolar Cell (LMC). Similar to bipolar cells in human visual pathways, LMCs process information by removing redundant information (Srinivasan et al., 1982). Surround processing from the nine (eight adjacent and one centre) photoreceptors are summed, scaled, and then subtractively inhibited the centre LMC signal, with the scaling set to pass a 30% DC spatial component. Prior to inhibition, the surround signal was passed through a first-order temporal low pass filter ($\tau=2\text{ms}$), as an estimation of the time for signal spread of the surround interactions. The recombined LMC signal is then temporally filtered with a relaxed first order high pass filter (10% DC, $\tau=40\text{ms}$). This signal was inverted to emulate hyperpolarisation to luminance increments.

3.2.3.4 Rectifying Transient Cell

For the version of the model as presented in this paper, any sustaining component is removed by passing the signal through another high pass filter ($\tau=40\text{ms}$). The ‘on’ and ‘off’ phases are then separated into distinct channels via half-wave rectification (with one signal then inverted). Each channel is temporally processed via a fast adaptive mechanism. An ‘adaptation state’ nonlinear signal F , as represented by equation 2, is created via a low pass filter that can transition between two time constants. If the original ‘pass-through’ signal (X) was increasing, the adaptation state time constant was 1ms (τ_1), if the pass-through signal was decreasing then the time constant was set at 100ms (τ_2). This allowed for representation of a fast depolarisation (increasing signal) but slow repolarisation (decreasing signal) cellular effect.

$$\begin{aligned} d/dt\{F\} &= (X - F) / \tau_1 \quad (X - F \geq 0) \\ &= (X - F) / \tau_2 \quad (X - F < 0) \end{aligned} \quad (2)$$

In each of the ‘on’ and ‘off’ channels, this adaptation state subtractively inhibited the pass-through signal. The outcome of this was a fast adaptive mechanism as seen in response properties of the RTC.

A spatial centre-surround antagonism was applied to each of the channels, i.e. an ‘on’ surround inhibiting the ‘on’ centre and an ‘off’ surround inhibiting the ‘off’ centre. These surround signals included a neural delay ($\tau=2\text{ms}$). This inhibitory surround interaction should not reduce the centre output below zero so a further half-wave rectification was applied.

Finally, the detector was tuned for sensitivity to dark targets by implementing either a second order or a linear recombination of the delayed ‘off’ channel (via low pass filtering, $\tau=25\text{ms}$) with the undelayed ‘on’ channel.

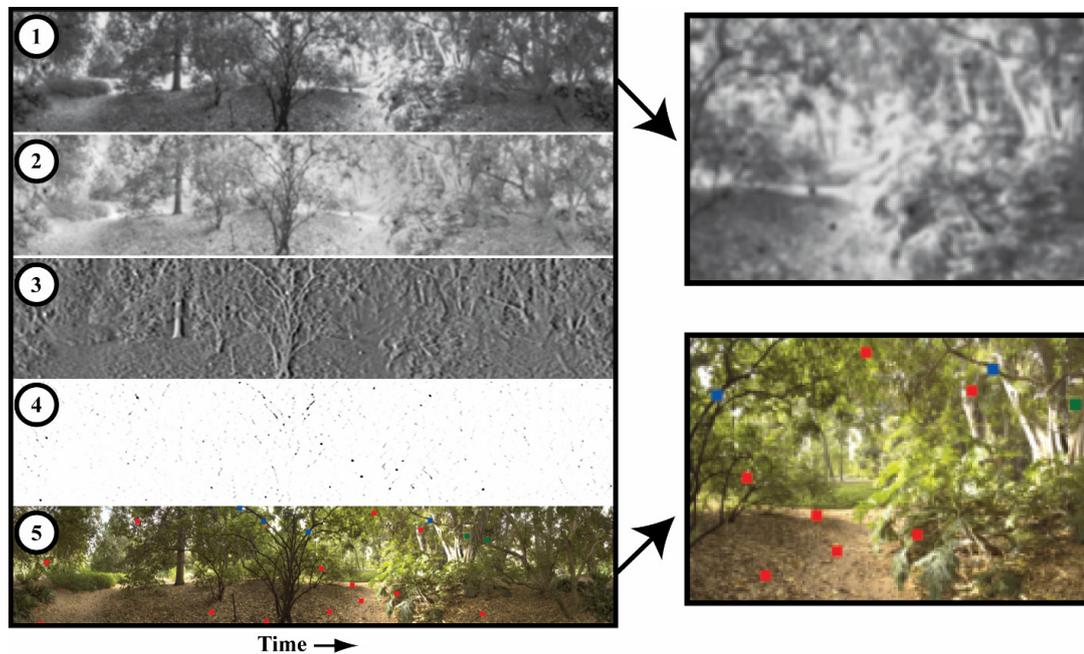


Figure 3-3. An example of model outputs. Panel 1 shows spatial blurring (inherent in fly optics) and the ‘green’ channel of an RGB image (spectral sensitivity of fly motion detection). Panel 2 is at the photoreceptor level, including a dynamic, compressive nonlinearity. Panel 3, spatiotemporal band pass filtering of the LMC stage. Panel 4 reveals eSTMD outputs and panel 5 shows a particular threshold of eSTMD output, superimposed on the original image (red = hits, blue = false, green = misses).

At each stage of processing, model outputs were examined and target position determined via cross-correlation of images ‘with’ and ‘without’ targets. Output responses were then binned in time, dependent on the panorama velocity (transformed back into a spatial domain). At any given processing stage, varying threshold levels of the outputs were examined and bins were categorised, with each either a response to a target (a hit), or alternatively a false positive response. Receiver Operating Characteristic (ROC) curves were then produced to indicate the success of target enhancement / non-target suppression at the various stages of the eSTMD processing.

3.2.4 Results

Fig. 3 shows the results from an example simulation of the model. The panels are a reconstruction from a column of outputs with responses over time, at the various output stages. The first panel reveals the difficulty of the target detection task when considering the low spatial resolution and single spectral sensitivity that serves as inputs to the higher levels of processing. Photoreceptor output shown in panel 2 applies a dynamic saturating nonlinearity, which is most beneficial when considering high dynamic range inputs, rather than the 8-bit representations (though input values were calculated with linear interpolation) used in this study. The LMC stage of panel 3 shows the effects of spatial and temporal band pass filtering, the parameters of which are based on physiological time constants. This stage is the first of two used in further ROC comparisons. Panel 4 is the final eSTMD multiplicative recombination, which includes the processing inherent in the RTC. In panel 5, we look at an example thresholded eSTMD output, categorised and superimposed on the original input image. Both the LMC and eSTMD stages are shown, for a variety of natural images, in the ROC curves of Fig. 4.

ROC curves lying in the upper left corner of the graph, i.e. with high hit rates and low false positives, reveal enhancement of target discrimination. The absolute number of false positives is dependent on bin size, however it is the relative change between curves that is significant in the target discrimination task. The input image ‘All’ is reconstructed from the average magnitude and average phase components of thirteen natural scenes and all of these images have a characteristic ‘natural scene’ power spectrum response of approximately $1/f$, with spatial frequency f (Field, 1987). During target distribution, some targets are scattered on to near equi-luminant backgrounds so that there is little or no basis for discrimination, and 100% hit rates are not to be expected. This upper constraint on target detection can most readily be seen in the ROC of the most cluttered scene, ‘Botanic’. In the image ‘Field’ there is little difference between ROC curves as the spatiotemporal band pass filtering has effectively discriminated the targets, due to the lack of other contrast features in the image.

In the other three images, the enhancement in target discrimination is evident by a shift of the ROC curves into the upper left corner.

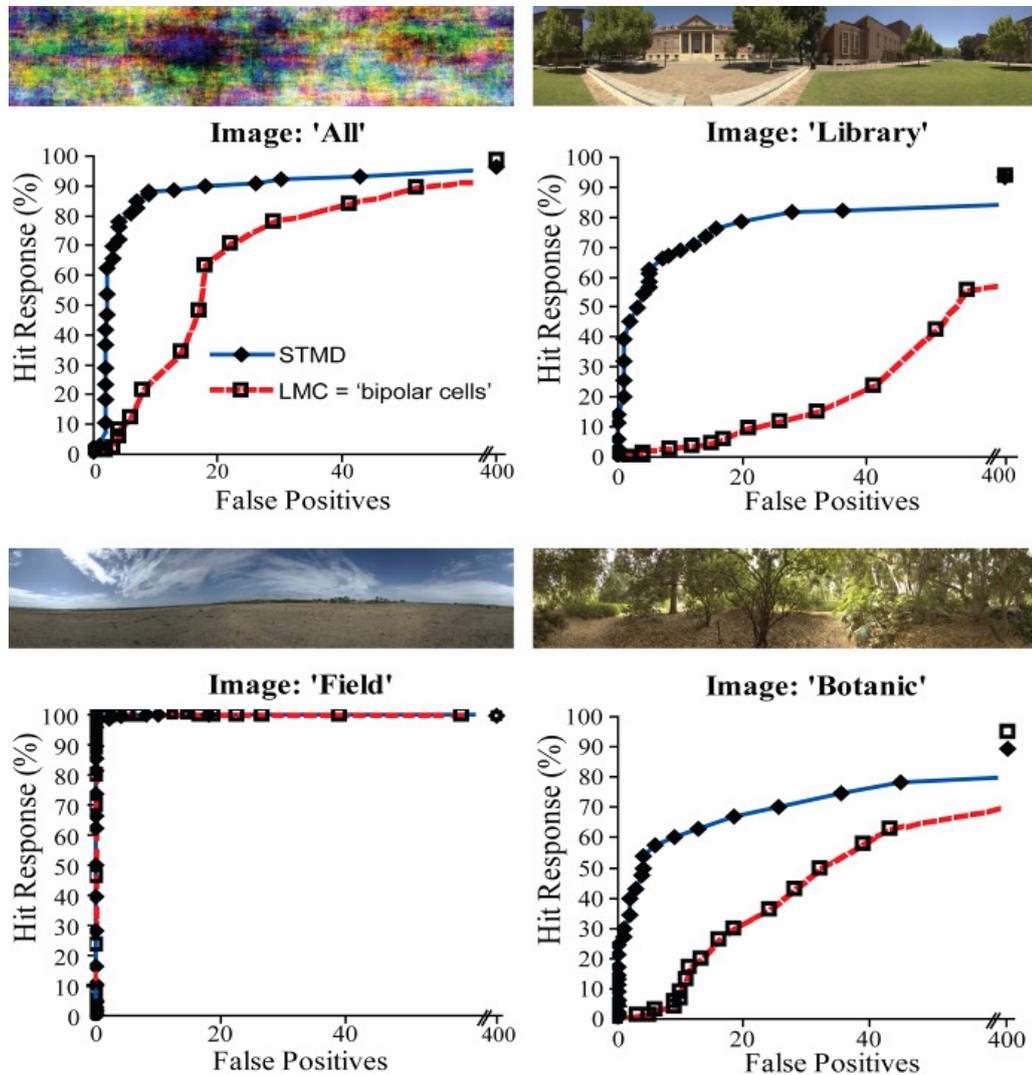


Figure 3-4. ROC curves for four panoramic images. Curves that lie in the upper left corner (high hits, low false positives) indicate processing that is more effective at discriminating targets.

3.2.5 Discussion

The results from these simulations suggest that within these four natural scene images, spatial statistics similar to that of small targets are sufficiently rare that a highly nonlinear filter that can exploit these. For a dark target, the requirement for an unadapted, or novel, ‘off’ phase provides a temporal distinctiveness to the target which is followed by an unadapted ‘on’ phase, inherent in the target itself. This, combined

with a requirement for these same aspects *to not be* present in the surround, provides us with small target selectivity, irrespective of the other characteristics of background. If this final parameter, the surround ‘on’ or ‘off’ channel inhibition, is relaxed, this model can be tuned for bar sensitivity, as seen in some types of insect neurons (O’Carroll, 1993).

Characteristic responses from this model match the neuronal STMD in several ways. Neuronal STMDs are tuned to velocity and this is the same in the model due to the temporal correlation of the ‘on’ and ‘off’ phases. For their small target selectivity, STMDs neurons are target height tuned, and this is replicated in this model due to the inhibitory surround mechanisms. Finally, some evidence suggests that small target selective neurons in hoverfly (Nordström et al., 2006) are superlinearly dependent on contrast, as is also seen from the multiplicative structure within this model. Generally, the model properties match those of the physiological correlate. However, a collection of these units, perhaps pooled, would be required to form the position invariant structure of neural STMDs. Direction selectivity, as seen in some of these neurons, could be formed in our modelling with an asymmetry in the inhibitory surrounds or could also emerge during the process of pooling eSTMD units, i.e. a directional spatial facilitation. A spatial facilitation of eSTMD outputs might further enhance target discrimination and is a focus of future modelling. This model highlights that RTC processing is a feasible mechanism for the detection of small moving targets and further physiological experimentation on STMD and RTC neurons will attempt to examine this possible role in greater detail.

A common approach to modelling figure-ground discrimination tasks, such as this small target detection problem, has focused on the use of relative motion cues. This involves a computationally intensive technique of contrasting motion components of figure versus ground. We have found in our modelling, inspired from a physiologically available unit of processing (the RTC), that a highly nonlinear ‘matched’ filter may be all that is required for the target discrimination task. Our approach is well suited to solving what was initially a perplexing issue, i.e. the manner in which the neuronal STMD still responded to a small target even when the target was matched to the velocity of the background. Further investigation will examine how matched temporal changes (with adaptive complexities) of other feature signatures may be useful in detection and tracking tasks.

3.2.6 Acknowledgement

This work was supported by grants from the US Air Force Office of Scientific Research (AFOSR), FA9550-04-1-0294, FA9550-04-010283.

Panoramic images courtesy of Russell Brinkworth.

CHAPTER 4:

TARGET DISCRIMINATION IN GENERAL STIMULUS CONDITIONS

4.1 CONTEXT

The previous modelling efforts had provided evidence for the feasibility of RTC-type processing in a model for target discrimination within visual clutter. Due to limitations in the length of the PLoS article (with the new physiological results), more simulations were left for further publications. In this next publication, we examine target detection under more general stimulus conditions.

We also took the opportunity to extend the modelling, into a more physiologically accurate realm.

1. The model inputs were changed to accept high dynamic range (HDR) values (therefore representing ‘real-world’ luminance and contrasts). This allowed for testing of target detection in simulations more akin to visual ecology.
2. A spatial transformation was implemented into the hexagonal domain to approximately represent the arrangement of facets in insect vision systems. The spatial arrangement is maintained throughout the visual processing pathway, so that higher order spatial interactions, deeper within the visual system, are between hexagonally arranged units.
3. The placement of targets was not aligned to the centre of image rows, thus making target detection more difficult.
4. A laboratory research interest is in the development of a biomimetic representation of the photoreceptor. We included this photoreceptor model as a front end to our own target detection modelling. We could now address how the addition of complex photoreceptor dynamics would aid in the target discrimination task in more realistic visual environments.

5. We searched the literature to develop a better representation of the 1st order interneuron (the LMC). By examining frequency response curves at various background luminance values we developed some adaptive temporal filtering characteristics, and also better represented spatial inhibitory surrounds.

With these model elaborations, we simulated the more general stimulus conditions. This included using a series of 18 HDR floating-point data sets (natural images) with realistic (and precise) representation of visual environments.

For this article we used a traditional quantification method for target discrimination. This involves developing receiver operating characteristic (ROC) curves and integrating under the curve to find a value for the ‘area under the ROC’ (AUROC).

We simulated the target detection mechanism (with pseudo-randomly distributed targets) across changing target sizes and panorama (and therefore target) velocities. Relative motion considerations have already been investigated in the previous PloS ONE (see 2.2) publication.

These simulations led us to determine that the model output depends on target velocity, more so than overall luminance or contrast conditions.

4.2 BIO-INSPIRED SMALL TARGET DISCRIMINATION IN HIGH DYNAMIC RANGE NATURAL SCENES

S. D. WIEDERMAN¹, R.S.A. BRINKWORTH¹, D.C. O'CARROLL¹

¹*Discipline of Physiology, School of Molecular and Biomedical Science, The University of Adelaide, Adelaide, Australia*

ACCEPTED: Wiederman, S.D., Brinkworth, R.S.A. and O'Carroll, D.C. (2008) Bio-inspired small target discrimination in high dynamic range natural scenes. Proceedings IEEE, Bio-inspired Computing: Theories and Applications (BIC-TA), in press.

4.2.1 Abstract

Flies have the capability to detect and track small moving objects, often against cluttered moving backgrounds. From both a physiological and engineering perspective, understanding this computational process is an intriguing challenge. We have developed a target detection model inspired from electrophysiological recordings of 'small target motion detector' neurons within the insect brain. Our numerical modeling represents the neural processing along a proposed pathway to this target-detecting neuron. We use high dynamic range, natural images, to represent 'real-world' luminance values that serve as inputs to a biomimetic representation of photoreceptor processing. Adaptive spatiotemporal high-pass filtering (1st-order interneurons) then shape the transient 'edge-like' responses, useful for feature discrimination. Nonlinear facilitation of independent 'on' and 'off' polarity channels (the rectifying, transient cells) allows for target discrimination from background, without the need for relative motion cues. We show that this form of feature discrimination works with targets embedded in a set of natural panoramic scenes that are animated to simulate rotation of the viewing platform. The model produces robust target discrimination across a biologically plausible range of target sizes and a range of velocities. Finally, the output of the model for small target motion detection is highly correlated to the velocity of the stimulus but not other background statistics, such as local brightness or contrast, which normally influence target detection tasks.

4.2.2 Introduction

The ability to detect and track small moving objects allows flies to engage in rapid chase pursuits (Collett and Land, 1975, Wehrhahn, 1979). The computation underlying this visual detection of moving targets is non-trivial, especially considering that they are often viewed against complex, moving backgrounds. A spatial acuity of $\sim 1^\circ$ is a further limiting constraint (Straw et al., 2006). This capability to perform real-time target detection and tracking is a humbling thought in view of the small size, weight and number of neurons within a typical insect brain (Strausfeld, 1976). This consideration has led researchers to study the behavioral aspects of small target detection, such as examining fly responses to moving objects within experimental flight arenas (Boeddeker et al., 2003).

A different research approach is via electrophysiological experimentation. Intracellular, in vivo recordings, from the optic lobes of the fly brain, reveal possible neural correlates to the target detection task. Such cell types have been encountered in a range of insects including; blowfly (Wachenfeld, 1994), dragonfly (O'Carroll, 1993) and fleshfly (Gilbert and Strausfeld, 1991). One such neuron class, within the hoverfly *Eristalis tenax*, referred to as the Small Target Motion Detector (STMD), has been examined and many characteristics described (Nordström and O'Carroll, 2006, Nordström et al., 2006, Barnett et al., 2007). The response of the STMD neuron is selectively tuned to small objects (Nordström and O'Carroll, 2006), sensitive to contrast, and position invariant within a section of the visual field (Nordström et al., 2006). The cell exhibits tuning for preferred velocities and some are direction selective. A subset of these neurons respond to the movement of small targets, even when the target velocity is matched to that of the moving background, thus ruling out relative motion cues as a defining factor in the neural computation (Nordström et al., 2006).

In order to derive models for this processing, we examine the higher-order properties of the STMD in conjunction with neuron types that are candidates for inclusion along this visual pathway. The photoreceptors are the only cell type that we may definitively assert are at the front end of any visual task. However, we may hypothesize roles for other neuron classes, such as the first-order interneurons, the Large Monopolar Cells (LMCs), as these are believed to be involved in motion processing (Coombe et al., 1989, Douglass and Strausfeld, 1995). The LMCs have adaptive temporal band-pass

characteristics and spatial antagonistic interactions (James, 1992, Srinivasan et al., 1982). Hence are well suited to shaping the transient, spatially localized responses required for target detection.

We have recorded from a class of neurons referred to as the Rectifying Transient Cell (RTC) (Wiederman et al., 2008), which exhibit a functionality seen within a variety of invertebrate visual neurons (Arnett, 1971, Jansonius and van Hateren, 1991, Osorio, 1987). These cells show independent adaptation to contrast increments versus contrast decrements and exhibit surround antagonistic interactions (Jansonius and Van Hateren, 1993a). Previously, we examined response characteristics of types of ‘on-off’ transient neurons, including the RTC, and have shown that they are well suited as an adaptive, matched spatiotemporal filter, selectively discriminating target from background within moving panoramic images (with or without relative motion considerations) (Wiederman et al., 2008).

Realistic target-background scenes have a much larger range of luminance, contrasts and spatiotemporal structure than can be captured using traditional 8-bit digital imaging technology. To analyze the effectiveness of our small target detection methods across these ‘real-world’ conditions, we must have suitable front end high dynamic range input imagery (Debevec and Malik, 1997) and a sound emulation of early visual processing (Mah et al., 2008a). Using both biomimetic and bio-inspired neural elements we have built a numerical model that effectively increases target discrimination across a range of natural scenes, and have shown that this modeling is robust across varied target sizes and velocities.

4.2.3 Methods (Modeling & Simulation)

4.2.3.1 Overview

Image preparation, model simulations and analysis tools were written and implemented in Matlab (Mathworks, Natick, USA). A model overview is shown in Figure 1. The values for model parameters were set with respect to the underlying biological system. Photoreceptor responses were based on a biomimetic model (Figure 2) (Brinkworth et al., 2006, van Hateren and Snippe, 2001) with parameters and elaborations derived via electrophysiological results from *Eristalis tenax* (Mah et al., 2005). The spatiotemporal dynamics of the first order interneurons, the LMCs, have been previously established for the blowfly (*Calliphora vicina*) (1995) and hoverfly, (*Eristalis tenax*) (1992). A bio-

inspired role for RTCs in the small target detecting pathway was based on parameters from the ‘on-off’ transient cells (*Calliphora vicina*) (1991) and those found from our own electrophysiological recordings of RTCs in blowflies (*Calliphora stygia*) (Wiederman et al., 2008). The model outputs were matched to those seen in the STMD neurons as described for *Eristalis tenax* (Barnett et al., 2007, Nordström et al., 2006, Nordström and O'Carroll, 2006).

We modeled an array of ESTMD (elementary STMD) subunits that could be pooled to form the position invariant receptive field, as seen in the physiological STMD (Nordström and O'Carroll, 2006). The model outputs defined values in a three dimensional (2D+t) space, at each level of visual processing; the photoreceptor, LMC and ESTMD.

4.2.3.2 Input Images

The high dynamic range (HDR) data set for this modeling was a series of eighteen natural images, captured on a Nikon D-70 camera. Photographs were taken at various locations providing images of different luminance, contrasts and spatial structure and have been previously published (Brinkworth and O'Carroll, 2007). The photographs have natural image statistics, with a $1/f^2$ relationship between spatial frequency and power (Field, 1987). Panoramas with a 360° horizontal by 72° vertical extent were created by stitching together a series of twelve overlapping images. Each image was photographed at a range of exposure levels and HDR values reconstructed via a calibration technique (Debevec and Malik, 1997). From this we obtained floating-point data sets (realistic representations of luminance) to serve as inputs for our modeling purposes. Figure 3 shows some examples of these natural scene images.

Panoramic yaw rotation was simulated by animating these data sets at varying velocities. For the target detection task, fifty black squares were pseudo-randomly distributed over the images, with a minimum spatial separation of 5° vertical and with no repeated target on any horizontal row. This was to ensure target independence (no spatiotemporal interactions between the targets at any stage of processing). During simulation of the panorama motion, the targets were fixed to the background, thus removing relative motion cues from the discrimination task.

Motion processing pathways within the fly visual system are known to be monochromatic, with a ‘green-blue’ spectral sensitivity (Srinivasan and Guy, 1990). We therefore used the green channel of our RGB data set as an approximation.

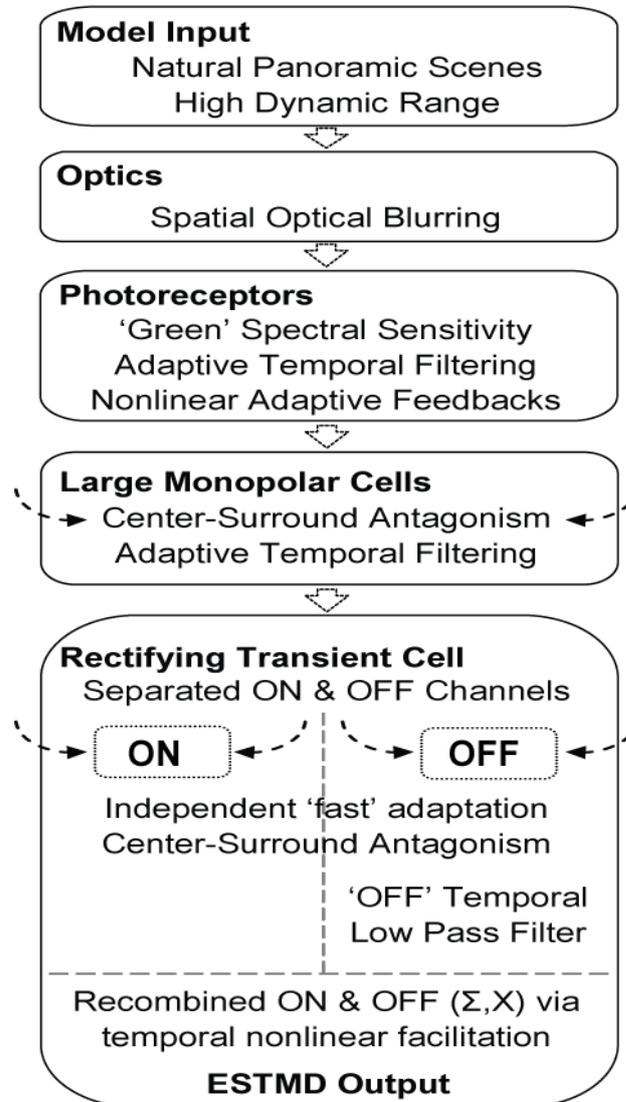


Figure 4-1. Overview of the ESTMD model. Motion of HDR panoramic imagery is simulated at varied velocities. These are spatially blurred to represent the fly optics. The green channel is retained to emulate the spectral sensitivity of the motion pathway. A complex photoreceptor model implements dynamic 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.

To emulate the spatial blur inherent in the optics of the fly visual system (Stavenga, 2003) we apply a Gaussian low-pass filter (full width at half maximum 1.4°) to the input images. We simulated a range of target sizes. While all targets had a defined luminance of 0 the process of blurring caused target luminance to vary, dependent on the immediate background, therefore the model responses to different targets are essentially to targets of varying effective contrasts.

The input images were spatially sampled at 1° in a hexagonal manner (Straw et al., 2006). This spatial organization was at all levels of processing (spatial interactions were between adjacent hexagonal ‘facets’ as seen in the fly visual system). The original target distributions were not centered on the sampled rows, ensuring the discrimination task was especially challenging. Original images had a resolution of 8000 x 1600 pixels and after the 1° hexagonal sampling produced 2D arrays of 360 x 62 model units. As the panoramic images were ‘rotated past’ the model input, at varied velocities and high emulated frame rates, linear sub-pixel interpolation in the spatial domain was used to ensure high resolution in the resultant temporal domain.

4.2.3.3 Photoreceptors

Blowfly photoreceptor dynamics have previously been modeled, incorporating individual photoreceptor gain control and saturating nonlinearities (2001). Elaborations to this model are a continued research focus (Mah et al., 2008b) and have been implemented in Matlab as the front-end element to this ESTMD model (Figure 2). The elaborated photoreceptor model (Brinkworth et al., 2006), includes variable gain control and dynamic low-pass filtering, with cutoff frequencies dependent on adaptation state. Followed by two divisive feedbacks (one linear, one exponential) representing short and longer term adaptations. Finally, a compressive, saturating nonlinearity is implemented by a Naka-Rushton transform.

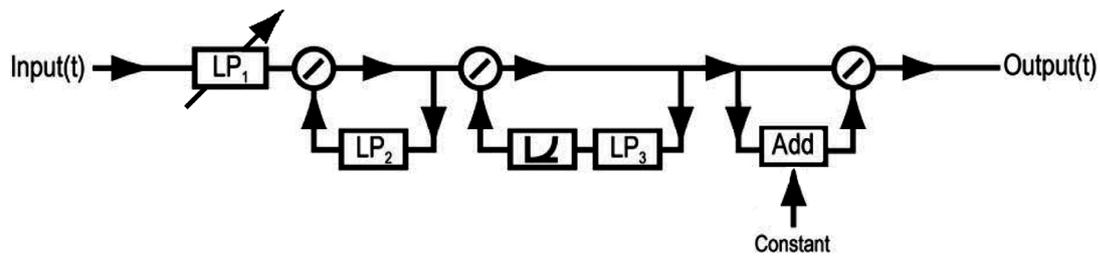


Figure 4-2. Complex Photoreceptor Model. Block diagram summarizing a biomimetic model of the photoreceptor, modified after Mah et al. (Mah et al., 2008a) and van Hateren & Snippe (van Hateren and Snippe, 2001). Photoreceptors allow for encoding a high dynamic range of real-world luminance values into the limited dynamic range of the neuron. This biomimetic model copies this process via a series of adaptive feedback mechanisms. The photoreceptor dynamics expand the possible range of inputs that can then be processed by the higher-order feature detection elements of the model. (LP1: feedforward, variable low-pass filter, LP2: linear, divisive feedback, LP3: exponential, longer term, divisive feedback, final stage: Naka-Rushton compressive nonlinearity)

4.2.3.4 Large Monopolar Cells

The LMC is the first order interneuron in the fly visual system and due to spatial and temporal antagonism (high-pass filtering) removes redundant information from correlations in visual scenes (Srinivasan et al., 1982) and are analogous to human bipolar cells. A more thorough examination reveals that the LMC alters 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 (Juusola et al., 1995). This spatial interaction (center-surround antagonism) was modeled in a feed forward manner with surround (nearest neighbor) photoreceptor signals summed, and temporally delayed, before subtractively inhibiting the central LMC. The LMC temporal dynamics were modeled with relaxed, variable high-pass filtering which incorporated a small DC component

4.2.3.5 Rectifying Transient Cells

The LMC output was fed into a Matlab version of our RTC model. A more thorough description of the RTC processing has been previously described (Wiederman et al., 2008). Briefly, transient ‘on’ and ‘off’ phases (from the LMC high-pass filtering) are separated via 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’ ($\tau = 3\text{ms}$) when channel input is increasing and ‘slow’ ($\tau = 70\text{ms}$) when decreasing). This adaptation state subtractively inhibits 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 subtractively inhibiting the ON centre channel, and similarly with the OFF channels. The resultant signal was then half-wave rectified, so that the surround does not inhibit the centre below a zero value (a nonlinearity seen in some spiking neurons).

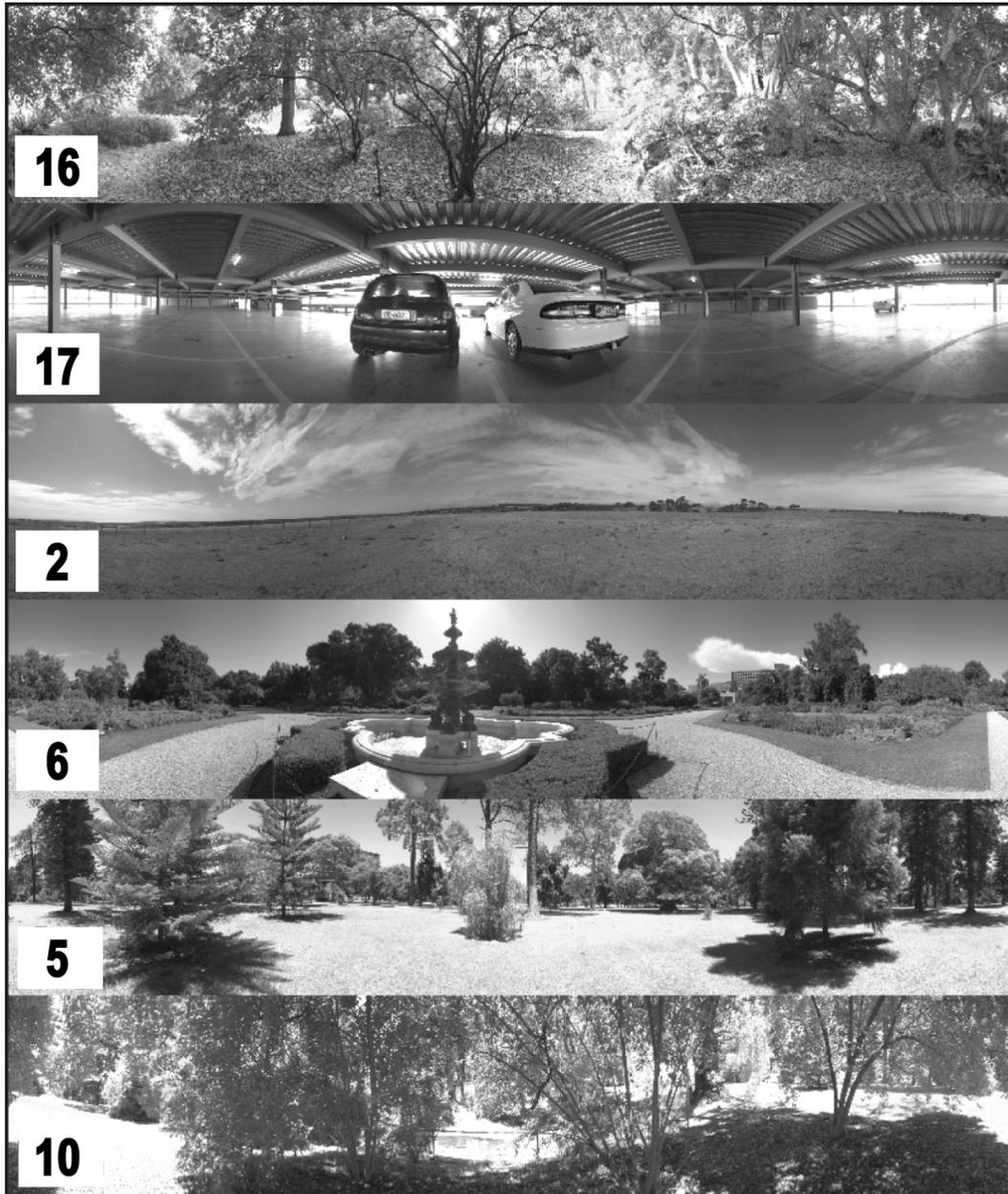


Figure 4-3. HDR Panoramic Images. Example panoramic images used as inputs for the target detection model. Images have been gamma corrected and reduced to 8-bits (for reproduction) and only the green channel (as used in the model) is shown. The original input images were 8000x1600 pixels and high dynamic range (HDR), meaning that there are no over or under saturated pixels and that the real-world luminance information was retained. The natural scene inputs range from an average luminance of 276 cd/m² to 17993 cd/m². Global image contrast values ranged from 0.73 to 4.6 (standard deviation/mean). All high dynamic range images are available upon request. Results for target detection in individual images can be seen in Figure 6 and matched by image number.

The two channels were then recombined by delaying the OFF channel (low-pass filter, $\tau = 25\text{ms}$) and multiplying this by the undelayed ON channel. This is similar to proposed correlation-type models for motion detection (Egelhaaf and Kern, 2002, Hassenstein and Reichardt, 1956), but the ‘delay and compare’ is with ‘off’ and ‘on’ transients from a single point in space rather than two spatially distinct points.

4.2.3.6 Simulation and Analysis

Simulations were run across a wide range of panorama velocities ($1^\circ/\text{s}$ - $1000^\circ/\text{s}$) on eighteen distinct HDR natural images, both with and without $1.4^\circ \times 1.4^\circ$ targets ($n=50$) inserted. Also simulated, at a near optimal velocity of $90^\circ/\text{s}$, three trials (50 targets each) repeated with 1.2° , 1.4° , 1.6° , and 1.8° target sizes. Taking differences between the ‘with’ and ‘without’ target images provides target location at each stage of the model output (luminance, photoreceptor, LMC and ESTMD) regardless of the non-linear temporal delays inherent in each stage of the processing. From the output stages we determined histograms of model values at the target locations in the ‘with target’ images ($n=50$) and histograms of model outputs in the ‘without target’ images ($n=360 \times 62 = 22320$). At any output threshold level, these histograms define the number of ‘hits’ (true positives, targets correctly classified as targets), ‘misses’ (false negatives, targets incorrectly classified as not targets) and ‘false alarms’ (false positives, non-targets incorrectly classified as targets). We used Receiver Operating Characteristic (ROC) curves (McNeil and Hanley, 1984) to represent target discrimination which plot hits versus false positives, across the varying model output thresholds. The ROC is usually quantified by examining the curve within the entire probability range, from 0 to 1 (hits from 0 to 50, false positives from 0 to 22320) however, for our purposes such a high false positive rate is not within a biologically plausible, or functionally relevant, range. We arbitrarily define our region of interest to not exceed 50 false positives, thus matching the total number of targets. By integrating under the ROC (up to 50 false positives) and normalizing, we quantify the effectiveness of target discrimination with a single value – the area under the ROC (AUROC). High AUROC values (close to 1) indicated the presence of a threshold producing good target hits and low false positives. Low AUROC values showed the discrimination between targets and non-targets was limited, hence the model performed poorly as a target detector.

4.2.4 Results & Discussion

An example of the simulation output is seen in Figure 4. Within the biological system, and emulated in this model, the adaptive mechanisms of the photoreceptor encode the high dynamic range of the luminance values within the restricted dynamic range of the neuron. This adaptation approximates a logarithmic encoding around an adaptation level (van Hateren and Snippe, 2001). The photoreceptors become temporally more responsive in the light adapted state, whilst in darker conditions require more temporal integration. Photoreceptor encoding is particularly responsive to decreased temporal contrast steps (Juusola et al., 1995), as encountered with a passing dark target.

The LMC stage reveals spatial and temporal high-pass filtering and therefore enhances edge-like contrast boundaries in both space and time. Target salience is improved from both photoreceptor and LMC processing, however few targets can be reliably separated from the background using a single threshold value (Wiederman et al., 2008). The ESTMD output improves the functionality as observed in the RTC model. In Figure 4 we have arbitrarily defined an output threshold and superimposed these results on the original blurred luminance image.

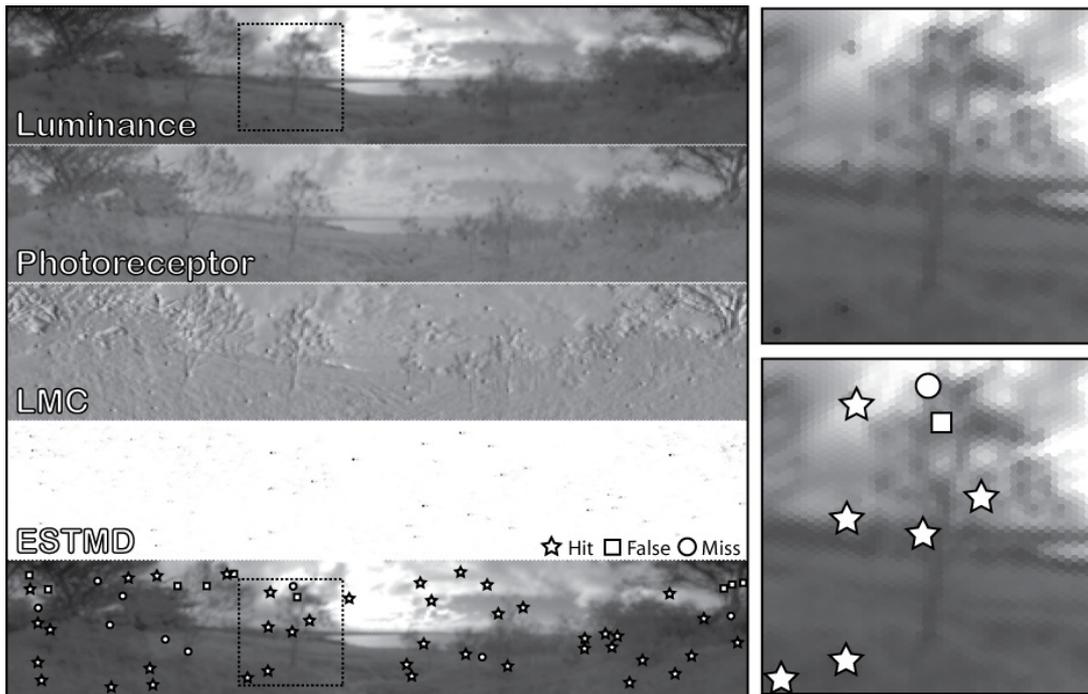


Figure 4-4. Model output stages. ‘Luminance’ shows spatial optical blurring and green spectral sensitivity. ‘Photoreceptor’ shows dynamic compressive adaptation, allowing for visual processing tasks, such as small target detection, to be viable within a large range of luminance conditions. ‘LMC’ portrays spatial and temporal high-pass filtering, enhancing the contrast boundaries that determine features within an image. ‘ESTMD’ output has suppressed temporal texture variations by signalling novel contrast changes, independently for ON and OFF channels. To be signalled by the ESTMD, this novelty must have a limited spatial extent, due to spatial antagonistic interactions. This processing is sensitive for target-type statistics, as shown in the bottom image where thresholded ESTMD values are superimposed on the luminance image. Insets show close-up views of the luminance both without (upper) and with (lower) superimposed model outputs (star = hit, square= false, circle = miss). Image number 11, panorama velocity 90°/s, target size 1.4°x1.4°. Despite the difficulty in detecting the targets in the original luminance image the full model correctly identified many of the targets with few false positives.

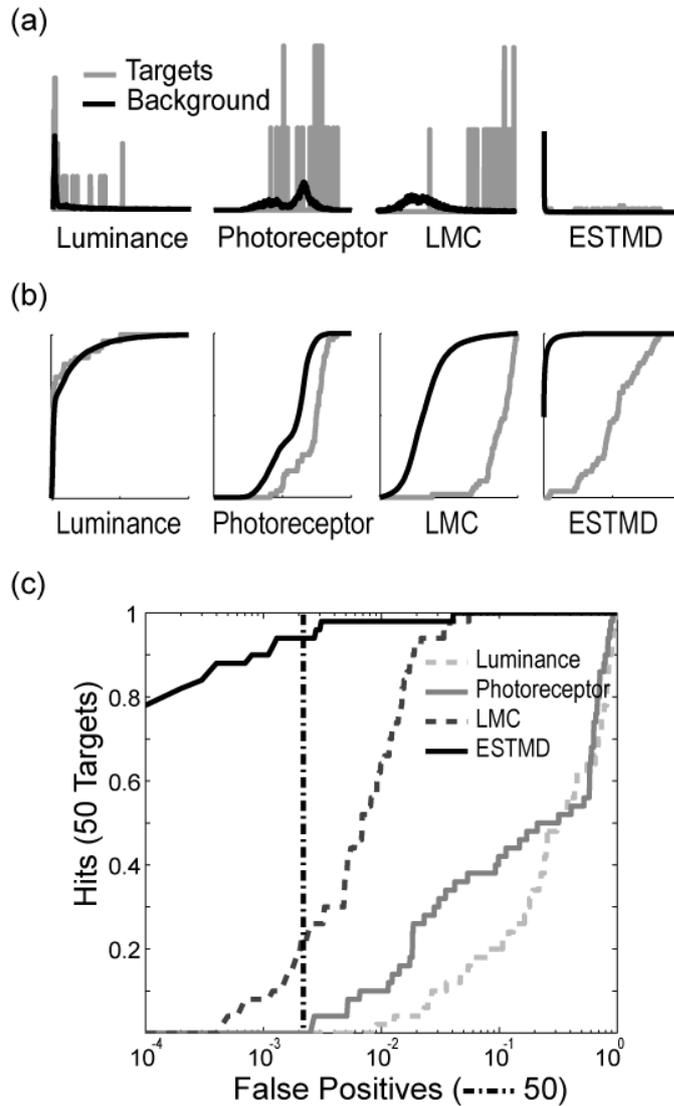


Figure 4-5. Receiver Operating Characteristic analysis. Histograms of output values for image number 11 at 90°/s and using targets of size $1.4^\circ \times 1.4^\circ$, as seen in Figure 4. (a) probability density functions (PDF) and in (b) cumulative distribution functions (CDF). Discrimination of target from background is revealed via non-overlapping PDFs and separable CDFs. (c) Receiver Operating Characteristic (ROC) curves are derived from the CDFs. ROC curves that lie in the upper left corner (high hits, low false positives) have more effectively discriminated targets from background. The area under the ROC is integrated to 50 false positives (dotted line), arbitrarily defined as a biologically plausible range and increases dramatically after ESTMD (elementary small target motion detector) processing. Luminance is the raw image values in high-dynamic range. Photoreceptor processing normalizes the images and compresses them to fit into a lower bandwidth. LMC processing is the equivalent of high-pass spatiotemporal filtering.

For analysis, we determine the effectiveness of discriminating target versus background via the ROC curve methodology. The stages of which are illustrated for a single example image in Figure 5. Based purely on the luminance input it is not possible to discriminate the targets from the background in any meaningful way as the histograms for the targets and the background overlap to such an extent. As the processing levels are applied (first the photoreceptor, then LMC and finally ESTMD) the cumulative distribution functions for the targets and background become increasingly separable with AUROCs rising from a 0 in the cases of both the luminance and photoreceptor processed images to 0.15 after processing by the LMCs and 0.79 after the ESTMD stage.

Figure 6 compares LMC and ESTMD AUROC values for varied target sizes across the entire series of natural images (at a simulated yaw velocity of $90^\circ/\text{s}$). Natural images are ordered with respect to their increasing RMS Contrast (global standard deviation divided by mean). The ability to discriminate targets is least in the low brightness and strong textures of image 16. At the other end of the scale is image 2, which is sparse and allows for the rare case of effective discrimination at the level of the LMC (see Figure 3 for images). Between these two images lies an assortment of natural conditions where the RTC processing has improved the AUROC values in comparison to the LMC stage. During the pseudo-random distribution, some targets are scattered onto backgrounds where the result is the loss of all ‘target-like’ defining characteristics. This is especially the case in highly textured sections of images (see Figure 4 inset for such an example). Therefore, in most cases, we do not expect AUROC values to approach unity.

In Figure 7, we show the improvement in AUROC between the ESTMD and LMC stages, for each of the images, across the tested velocity range (at a target size of 1.4°). The actual model outputs exhibit characteristic velocity tuning (as seen in the physiological STMD) and this is also observed in the AUROC analysis. Little or no improvement in target discrimination is observed outside the range $10 - 300^\circ/\text{s}$ with the optimal being $90^\circ/\text{s}$. This is a direct result of the RTC model tuning parameters only set to detect targets within this range. Below $10^\circ/\text{s}$ the target transients corresponding to the start and finish of the target are too temporally distinct to allow for a good correlation with the time constants used in the model. Above $300^\circ/\text{s}$ the low pass filtering of the

early visual processing blurs out the targets making them indistinguishable from the background.

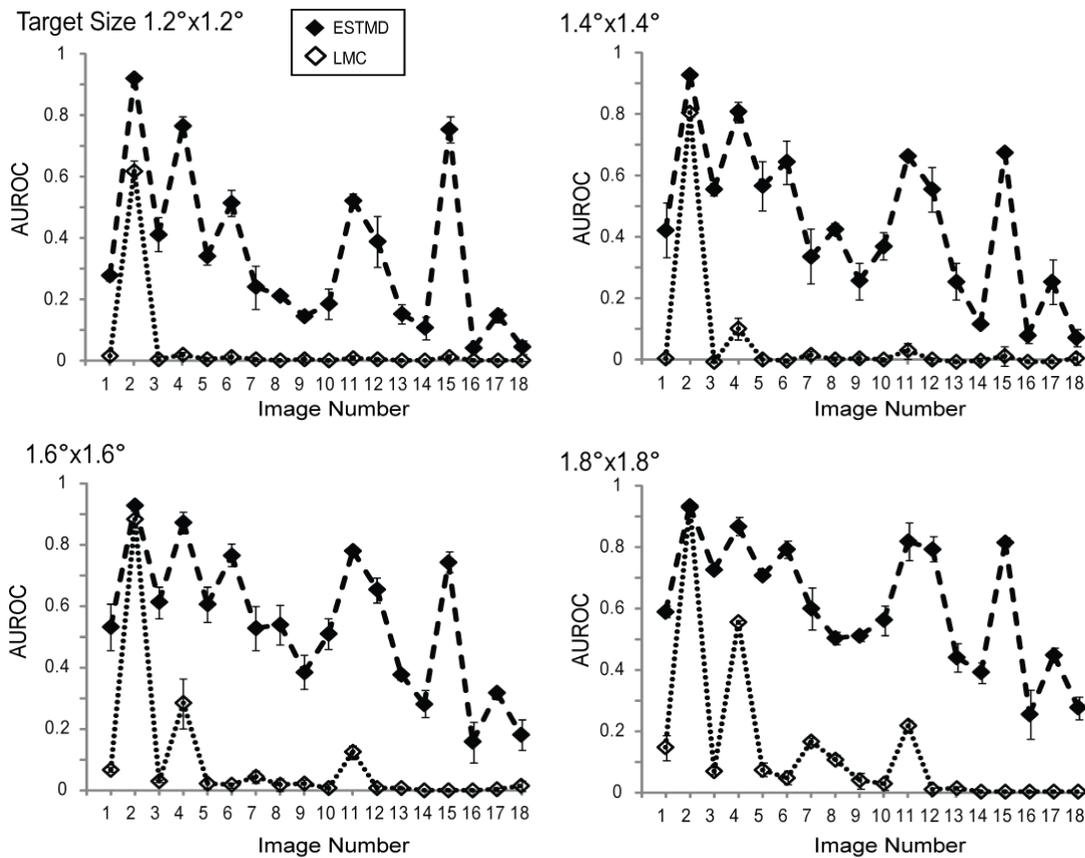


Figure 4-6. Target Size AUROC. AUROC values of four target sizes across 18 natural images (panorama velocity of 90°/s). An AUROC value of one would indicate the detection of all 50 targets before encountering any false positives. The elementary small target motion detection (ESTMD) processing enhances target discrimination from background, across varying scene luminance, contrasts and clutter. Comparison between closed symbols (LMC) and open symbols (ESTMD) show the AUROC improvement from the addition of RTC-type processing. This improvement is consistent across target sizes. Images are ordered with respect to increasing global contrast (standard deviation divided by the mean). Error bars indicate the AUROC standard deviation over repeated simulations (N=3) of each image, at a particular target size, with different pseudo-random target distributions. In all images and all target sizes the ESTMD processing produces better target discrimination than spatiotemporal high-pass filtering (LMC).

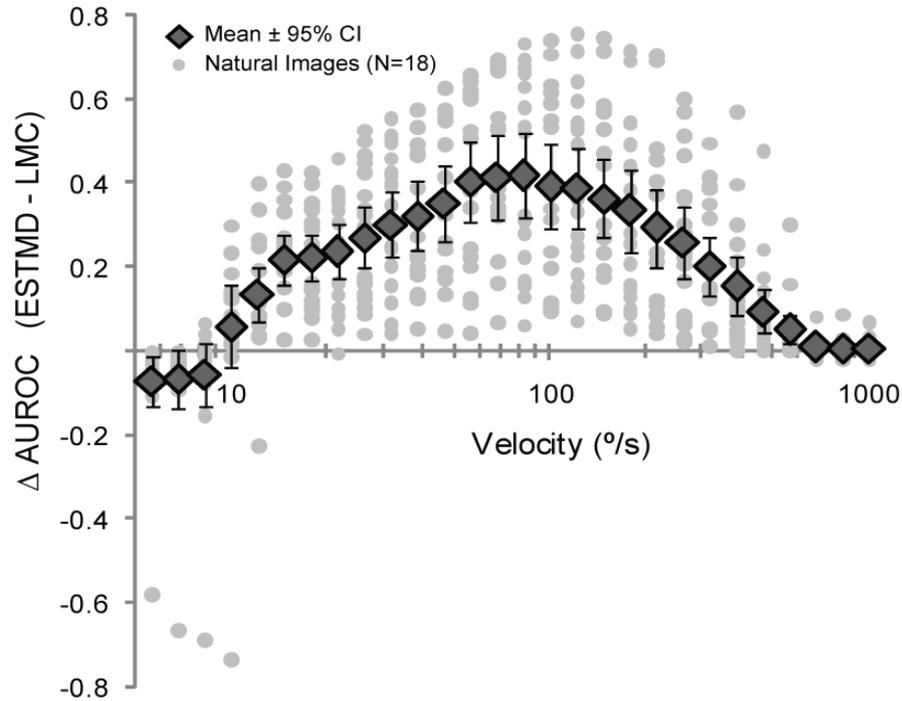


Figure 4-7. Panorama Velocity AUROC. The improvement in target discrimination (Δ AUROC) from spatiotemporal high-pass filtering alone (LMC) to the full elementary small target motion detection (ESTMD) model, shows consistency across panorama images and velocities (target size used $1.4^\circ \times 1.4^\circ$). As the delayed OFF channel is recombined with the ON channel, via a multiplicative nonlinearity, the model responses and final AUROC exhibit a velocity tuning curve indicative of correlative motion detection models. The peak velocity of this curve is dependent on the time constant of the OFF delay filter. The mean Δ AUROC is shown with estimated 95% confidence levels ($1.96 \times \text{SEM}$, $N=18$).

Finally, we examined the ESTMD responses of $1.6^\circ \times 1.6^\circ$ targets ($N=50$), across the 18 scenes, at 14 different velocities. At each target location, we look at the local luminance and contrast values ($5^\circ \times 5^\circ$) of the background, in order to derive coefficients of determination (R^2) between the image properties (local luminance, local contrast and panorama velocity) and model ESTMD responses. The results of which are shown in Figure 8. Local luminance is on a logarithmic scale (approximating photoreceptor encoding) and panorama velocity is on a log-log scale (represents the power function inherent in the multiplicative nonlinearity). We see that the ESTMD output is most strongly determined by target velocity, as would be expected in a motion energy detection mechanism, with little correlation between target responses and local image statistics. Scaling of the overall motion-dependent excitability of this system (and therefore optimal discrimination threshold value), as might be induced by global ego-motion, is a focus of future research in our target discrimination modeling.

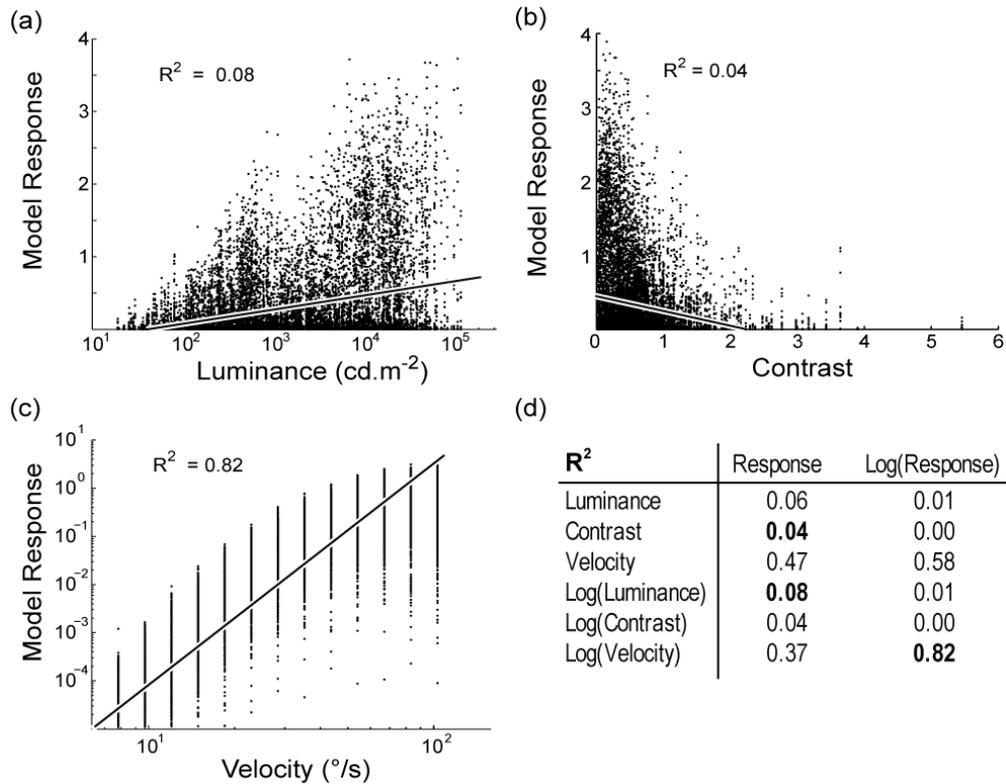


Figure 4-8. Model Response correlations. (a) Elementary small target motion detection (ESTMD) responses (at target locations) are plotted against their respective mean local background luminance ($5^\circ \times 5^\circ$). As luminance conditions increase there is only a small increase in the responsiveness of the ESTMD model. (b) As local contrast of the background increases, ESTMD responses at target locations slightly decrease. High contrast conditions reveal ‘cluttered’ regions with limited suppression of target responsiveness. (c) The ESTMD responses at target locations are well correlated to the velocity of the panorama image (and therefore velocity of the target), within the increasing range of the velocity tuning curve. This is due to the correlation-like structure of the target detection mechanism. Also note that the responses flatten out after about $70^\circ/\text{s}$ as the model leaves the linear coding range. This shows that the model outputs are more strongly influenced by the motion of the target than the local image statistics around the target. (d) The coefficient of determination between the model responses at target locations and the stimulus variables (and their respective log transforms). Conditions in bold, corresponding to the largest R^2 values, are graphed. Target size used was $1.6^\circ \times 1.6^\circ$.

4.2.5 Conclusion

Our previous research showed that target discrimination based on RTC-type processing was feasible in response to 8-bit images, at a single target size and single panorama velocity. This new work has extended the modeling into three dimensions (2D+t) and integrated complex early visual processing components (biomimetic photoreceptors and LMCs), with hexagonal sampling and unconstrained target positions. In this study, we show that our results are robust over different target sizes and a large velocity range. By

using HDR input imagery we also show that the small target detection modeling works in a range of real-world luminance and contrast conditions.

From both the biological and engineering perspective, we have seen that even without relative motion cues, adaptive mechanisms can extract temporal feature signatures (as seen in the physiological correlate). These computations are less intensive than traditional motion segregation methods and this approach may be useful in further detection and tracking tasks, especially with respect to more complex object detection.

4.2.6 Acknowledgements

This research was supported by the US Air Force Office of Scientific Research (FA 9550-04-1-0294) and from the Australian Research Council in the form of a Linkage Grant (LP0667744) and Post Doctoral Research Fellowship (RSAB).

CHAPTER 5:

ELABORATED MODELS FOR TARGET DETECTION

5.1 CONTEXT

The previous publications examined small target detection within a physiological framework, using biomimetic or bio-inspired reasoning to design all aspects of the model. In this chapter we take a different approach, examining extensions to the model that provide a useful insight from an engineering perspective. The presence of these elaborations, within the physiological context, can be investigated in future experimentation.

The publications consider two model elaborations:

1. ESTMD inhibition from near-local EMD derived motion energy. In this form the ‘feature correlator’ is antagonised by a more traditional ‘motion correlator’. By varying the amount of inhibition we determine an optimal amount that maximises the value of target discrimination (AUROC).
2. ESTMD re-scaling from a wide-field estimation of velocity. Due to the correlation mechanism, the distribution of ESTMD values (target and background) will shift depending on the panorama velocity (whilst maintaining the same target discrimination). Therefore, to have a static optimal threshold the output values need to be normalised by the value of the panorama velocity. That is, if the feature correlator is normalised by the amount of ego-motion the result is a static optimal threshold for target discrimination. The value of the panorama velocity is determined from the model (from wide-field integration of EMDs).

The following publication describes these model elaborations which are particularly useful when considering the implementation of the feature correlator on moving platforms.

5.2 BIO-INSPIRED TARGET DETECTION IN NATURAL SCENES: OPTIMAL THRESHOLDS AND EGO-MOTION

S. D. WIEDERMAN¹, R.S.A. BRINKWORTH¹, D.C. O'CARROLL¹

¹*Discipline of Physiology, School of Molecular and Biomedical Science, The University of Adelaide, Adelaide, Australia*

PUBLISHED: Wiederman, S.D., Brinkworth, R.S.A. and O'Carroll, D.C. (2008) Bio-inspired target detection in natural scenes: optimal thresholds and ego-motion, in *Biosensing, Proceedings of the SPIE Vol.7035, 70350Z*

5.2.1 Abstract

We have developed a numerical model of Small Target Motion Detector neurons, bio-inspired from electrophysiological experiments in the fly brain. These neurons respond selectively to small moving features within complex moving surrounds. Interestingly, these cells still respond robustly when the targets are embedded in the background, without relative motion cues. This model contains representations of neural elements along a proposed pathway to the target-detecting neuron and the resultant processing enhances target discrimination in moving scenes. The model encodes high dynamic range luminance values from natural images (via adaptive photoreceptor encoding) and then shapes the transient signals required for target discrimination (via adaptive spatiotemporal high-pass filtering). Following this, a model for Rectifying Transient Cells implements a nonlinear facilitation between rapidly adapting, and independent polarity contrast channels (an 'on' and an 'off' pathway) each with center-surround antagonism. The recombination of the channels results in increased discrimination of small targets, of approximately the size of a single pixel, without the need for relative motion cues. This method of feature discrimination contrasts with traditional target and background motion-field computations. We improve the target-detecting output with inhibition from correlation-type motion detectors, using a form of antagonism between our feature correlator and the more typical motion correlator. We also observe that a

changing optimal threshold is highly correlated to the value of observer ego-motion. We present an elaborated target detection model that allows for implementation of a static optimal threshold, by scaling the target discrimination mechanism with a model-derived velocity estimation of ego-motion.

5.2.2 Introduction

A fly, as it pursues another fly, in acrobatic and rapid movements, is performing an intriguing and challenging task. It must detect the other ‘small target’ whilst both are moving at high velocities, often against complex moving backgrounds and in varied luminance and contrast conditions (Collett and Land, 1975, Wehrhahn, 1979). The fly has the capability to perform this task with only a million neurons, and a brain about the size of a grain of rice (Strausfeld, 1976). A humbling thought, considering engineering based applications of target detection and tracking. We record, via electrophysiological techniques, from neurons in the insect brain referred to as ‘small target motion detectors’ (STMD) (Nordström and O’Carroll, 2006, Nordström et al., 2006, Barnett et al., 2007). This neuron is size selective, and tuned to the velocity of small targets that move within a region of the visual field (Nordström and O’Carroll, 2006). They are contrast sensitive and some are direction-selective. The neurons respond to small targets even in the presence of complex, moving backgrounds. Interestingly, a subset of STMD neurons respond robustly, even when the velocity of the target and the background are matched (Nordström et al., 2006).

There are many engineering based methods to emulate this capability of salient feature detection. Optical flow calculations are used to determine the motion of an object and/or surround and are based on a variety of mathematical approaches (Barron et al., 1994, Beauchemin and Barron, 1995). Salient object detection, with and without background motion and overall ego-motion considerations, have a variety of proposed solutions (Choi et al., 2006, Tang et al., 2007, Wixson, 2000). In biological modeling of motion processing, the mathematical approaches are purposefully limited to those proposed to relate to the underlying physiological system. Hence, attention is focused on correlation-type motion detection (Hassenstein and Reichardt, 1956) (or equivalent ‘motion energy’ models (Adelson and Bergen, 1985)) as there is a good deal of physiological evidence to support this as a biological motion framework (Haag et al., 2004).

The first element along the neural pathway to the STMD is the photoreceptor, whose adaptive processing allow for the encoding of a large range of luminance values, in an approximately logarithmic manner. More detailed modeling of the photoreceptor incorporates a series of complex feed-forward and feed-back interactions (Brinkworth et al., 2006, Mah et al., 2005, van Hateren and Snippe, 2001). Electrophysiological experiments have also defined parameters for the dynamic temporal filtering characteristics of the model photoreceptors (Mah et al., 2005).

After the photoreceptors, the 1st-order interneurons, the Large Monopolar Cells (LMC) are believed to be on the motion processing pathway (Coombe et al., 1989, Douglass and Strausfeld, 2003, Keller, 2002) and their transient high-pass filtering characteristics (in both space and time) are well suited as inputs to typical elementary motion detector (EMD) models (Brinkworth and O'Carroll, 2007, Harris et al., 1999, Shoemaker et al., 2005). EMD modeling is usually based on elaborated versions of the Reichardt correlator (correlation-type motion detection) (Hassenstein and Reichardt, 1956), where signals indicate motion using a 'delay-and-compare' between two distinct spatial sampling points. One of the most widely studied sensory neurons in the optic lobe of the fly – the lobula plate tangential cells (LPTC), can encode ego-motion (Hausen, 1993) and there is strong evidence that they are derived from an EMD-type framework (review, (Egelhaaf and Borst, 1993)). LPTCs are wide-field neurons proposed to spatially integrate over many EMD subunits (in a nonlinear manner) (Borst et al., 1995). This is not limited to insects, as there is also evidence for EMD-derived motion processing in mammals, including humans (Clifford et al., 1997, Clifford and Ibbotson, 2002). Traditionally, EMD models are dependent on contrast and spatial statistics, however, in the latest modeling approaches, nonlinear elaborations allow for reliable velocity estimation (Brinkworth and O'Carroll, 2007) due to contrast and motion adaptive mechanisms and the typical $1/f^2$ power spectrum of the spatial statistics in natural scenes (Dror et al., 2000). Figure-ground discrimination, such as in the small target detecting task, could be implemented via a local EMD subunit, that is inhibited from a wide-field LPTC (Egelhaaf, 1985c). However, this form of relative motion detection was not observed in responses from at least a subset of STMD neurons (Nordström et al., 2006).

We have previously proposed a mechanism for target discrimination bio-inspired from the ‘on-off’ unit in the fly brain (Jansonius and van Hateren, 1991) and from recordings of the Rectifying Transient Cell (RTC) in the fly medulla (Wiederman et al., 2008). We model the RTC to take LMC input and half-wave rectify the signal into separated ‘on’ and ‘off’ channels. These channels adapt independently and include strong center-surround antagonistic interactions. Following this, a correlation type mechanism, but with opposite contrast phases (a delayed ‘off’ phase, with an undelayed ‘on’ phase) formed the basis of a dark target sensitive ESTMD (elementary) model. This processing would be integrated over space (similar to the LPTC) to form the position invariant receptive field of the STMD.

The ESTMD model is a highly nonlinear matched filter for the spatiotemporal profile of a small moving target and can improve target discrimination with or without relative motion cues (Wiederman et al., 2008). In this manner, it matches the physiological responses of the STMD neuron. Importantly though, the effect of wide-field interactions and relative motion cues is a known component of insect visual processing. Behaviorally, bees use motion parallax for figure-ground discrimination (Srinivasan et al., 1990a). In the fly, both small-field motion and their respective inhibition from wide-field counterparts, has been extensively studied (even to the extent of ablating the inhibitory neuron) (Egelhaaf, 1985c, Warzecha et al., 1993). Finally, dragonflies use motion camouflage (minimizing relative motion cues) to disguise themselves during aerial maneuvers (Mizutani et al., 2003).

We therefore consider higher-order interactions beyond our ESTMD modeling, and the effect of an inhibitory EMD-derived component on our feature detection mechanism. We examine the effects of this interaction at two spatial scales, thus forming two distinct elaborations to the model. At a smaller spatial scale, we examine the effects of ‘near-local’ EMD inhibition, a type of antagonism between the two correlator forms. At a larger spatial scale, we examine how inhibition from a wide-field neuron (velocity estimator) can scale the ESTMD output to account for ego-motion and allow for a static optimal threshold of target discrimination.

5.2.3 Modeling, Simulation & Analysis

5.2.3.1 Model Overview

We implemented the ESTMD model and analysis tools in MATLAB (Mathworks, Natick, USA). Figure 1 shows an overview of the ESTMD model, which is described in the following sections. This present research details two extensions to an established modeling effort (1) ESTMD with near-local EMD inhibition and (2) ESTMD with ego-motion rescaling. We provide outlines of the relevant stages of the basic model for clarity, however, detailed descriptions of this processing are found in previous publications. The use of panoramic HDR input imagery to serve as ‘real-world luminance’ data sets for vision models is described in Brinkworth and O’Carroll (Brinkworth and O’Carroll, 2007). The biomimetic representation of photoreceptors (Brinkworth et al., 2008, Mah et al., 2008a, van Hateren and Snippe, 2001), RTCs (Wiederman et al., 2007) and EMDs (Brinkworth and O’Carroll, 2007, Hassenstein and Reichardt, 1956, Shoemaker et al., 2005) forms the computational elements for the elaborations as described in this research.

5.2.3.2 Bio-inspired & Biomimetic

Model parameters are set with respect to the underlying physiological system, or their best estimates. The photoreceptor model is based on electrophysiological results from the hoverfly (*Eristalis tenax*) (Mah et al., 2008a). The LMC, has spatiotemporal filtering characteristics described for the blowfly (*Calliphora vicina*) (Juusola et al., 1995) and *Eristalis tenax* (James, 1992). The RTC is a cell class we propose is well-suited for feature discrimination (Wiederman et al., 2007) and this functionality is based on our own recordings from blowflies (*Calliphora stygia*) (Wiederman et al., 2008) and previous recordings from ‘on-off’ transient cells in *Calliphora vicina* (Jansonius and van Hateren, 1991). Parameters for wide-field motion selective units are derived from electrophysiological studies of *Eristalis tenax* (Nordström et al., 2008). We matched final model outputs to the properties as described in the STMD literature (*Eristalis tenax*). (Barnett et al., 2007, Nordström et al., 2006, Nordström and O’Carroll, 2006).

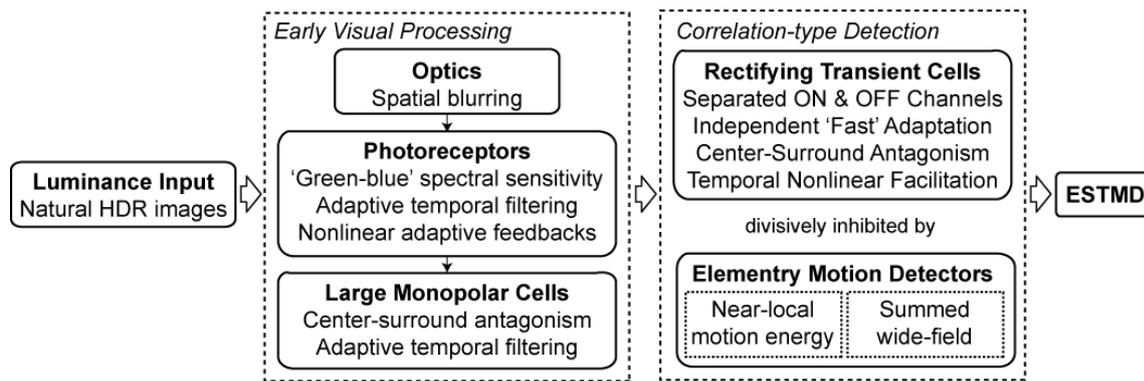


Figure 5-1. Elaborated Model Overview. High dynamic range (HDR) panoramic images, captured from natural scenes are animated at varying velocities, simulating ego-motion (yaw) of the viewing platform. We spatially blur the images to emulate the optics of the fly. Naturally occurring large luminance variations are encoded within the limited dynamic range of the photoreceptor via two ‘adaptive’ feedback mechanisms operating on different timescales. The photoreceptors also dynamically adjust temporal filtering characteristics based on the local conditions of the individual pixels. The Large Monopolar Cells high-pass filter the signal (relaxed), in both space and time, with adaptive filtering characteristics also dynamically set via the adaptation state. These transient signals serve as inputs to both correlation-type Elementary Motion Detectors (EMDs) and the Rectifying Transient Cells (RTCs). EMDs use a spatially separated ‘delay and compare’ facilitation, signaling the motion of ‘like’ contrast changes (a Reichardt correlator). They can be spatially integrated (wide-field) to provide an estimation of global velocity. The RTCs correlate ‘unlike’ contrast changes (‘on’ and ‘off’) at a single location to indicate the passing of feature boundaries (bar, target). Facilitation between delayed ‘off’ and undelayed ‘on’ transients provides dark feature sensitivity, whilst center-surround antagonism provides small target selectivity. Interaction between the two correlation types supports two forms of elaborations to the modeling. Near-local EMD inhibition of the RTC, decreases false positives in highly textured environments, at the cost of missed ‘hits’, however, the overall result is an improvement in target discrimination, especially in more challenging scenes. A second model elaboration is inhibition from the wide-field velocity estimation, which allows a static optimal threshold to be set for target discrimination that is robust across the entire velocity range.

5.2.3.3 Input Images

To serve as model inputs we use a series of 18 high dynamic range (HDR) natural scene images each with varying luminance, contrast and spatial structure (Brinkworth and O’Carroll, 2007). By capturing these images at a range of exposure levels, HDR values could be reconstructed (Debevec and Malik, 1997). The images extended 360° horizontal (panorama) and 72° vertical with a resolution of 8000 x 1600 pixels. Each of the images exhibited a $1/f^2$ relationship between spatial frequency and power, typically of natural scenes (Field, 1987). To determine the effectiveness of target discrimination we also replicate this set of 18 images but with 50 pseudo-randomly distributed targets (black squares, $1.6^\circ \times 1.6^\circ$).

We retain only the ‘green’ channel of the RGB image, approximating the green-blue spectral sensitivity of fly photoreceptors involved in the motion processing mechanism (Srinivasan and Guy, 1990). A spatial low-pass filter is applied (Gaussian with full width at half maximum of 1.4°) to replicate the optical blur of the compound eye (Stavenga, 2003). The image values are sampled in a hexagonal manner (like the facets of the fly eye), at a 1° spatial separation (Straw et al., 2006). With such limited spatial sampling resolution the targets to be detected are only 2.5 times the size of the area of a single pixel. In addition, due to the spatial blurring, the target luminance and contrast are varied across the distribution.

We simulated ‘yaw’ ego-motion (self-motion) by animating the panorama, rotating it past the model inputs at varying velocities (using sub-pixel interpolation in the spatial domain to ensure a high temporal resolution, minimum 1kHz). This sampling provided a spatiotemporal (2D+t) data set (360 by 62 model subunits).

During the target distribution, we maintained a minimum vertical separation of 5° , with no two targets on a single row (as the panorama was rotated horizontally). This ensured there were limited spatiotemporal interactions between the targets. Importantly, during the model simulations the distributed targets are embedded (fixed) to the background, removing any relative motion cues. Robustness of the target discrimination mechanism under varying relative motion conditions has previously been tested (Wiederman et al., 2008).

5.2.3.4 Early Visual Processing

Each element in the 2D hexagonal array was processed through a complex photoreceptor model (Mah et al., 2008a). This published model includes a variable feed-forward gain control and adaptive low-pass filtering. Both of these were set via an adaptation state (a low-pass filtered version of the luminance signal). For example, if luminance conditions are low, the gain is increased and corner frequency of the filter decreased (photoreceptor becomes slower and more integrative). Following this, two divisive feedbacks (via low-pass filters) representing two forms of adaptation, one a shorter term (a linear feedback) and the other longer term (an exponential feedback). Finally, we model a compressive ‘saturating’ nonlinearity with a Naka-Rushton

transform. With this processing, the photoreceptors could encode large luminance ranges as encountered in natural conditions.

The next stage of processing was via the LMCs. Similar to human bipolar cells, these neurons remove redundant information from correlations in the scene in both space and time (Srinivasan et al., 1982). We modeled this with variable high-pass filtering (again dependent on adaptation state), including a small DC component. As overall luminance conditions increase, the LMC becomes less sustaining and more transient in nature, adjusting their filtering characteristics for signal to noise considerations. The spatial high-pass filtering is typical feed-forward center-surround antagonism with nearest neighbor units summed, delayed, and then subtractively inhibiting the center signal. This type of processing is ideal for forming the transient edge-like signals (in both space and time) that are well suited for the further correlation-type processing.

5.2.3.5 Rectifying Transient Cells

Following the LMC we include a model for the RTC, our proposed intermediate stage neuron in the small target-detecting pathway (Wiederman et al., 2007). The signal is high-pass filtered, removing any DC component. We half-wave rectify the transient signal into two components – an ON channel and an OFF channel (inverted). We create an adaptation state for each channel via a nonlinear low-pass filter with a variable time constant. If the input signal is increasing, the time constant is set low (response is fast) or if the input signal is decreasing the time constant is set high (response is slow). Physiologically, this represents a neural response with fast depolarization and slow repolarization. This adaptation state signal subtractively inhibits the unaltered passed-through input signal. The result of this mechanism is one that suppresses rapid variations in the signal, only allowing a novel contrast change, or a large contrast change that can break-through the inhibited response. We apply this to each ON and OFF channel independently. Following this, we have a further center-surround antagonism between ‘like’ channels (subtractively inhibiting each other). Finally, the two are recombined by delaying the OFF and multiplying this with the undelayed ON channel in a correlation-type manner (but from a single point in space).

For maximum responsiveness, we require a temporally novel (unadapted) ‘off’ contrast boundary within a limited spatial extent shortly thereafter followed by a temporally

novel ‘on’ contrast boundary, again within a limited spatial extent. This is a highly nonlinear matched spatiotemporal filter for the detection of small dark moving targets.

5.2.3.6 Elementary Motion Detectors

The LMCs also serve as inputs to a more traditional type of EMD, based on the Reichardt correlator (Hassenstein and Reichardt, 1956). The EMD extracts motion information by sampling from two distinct spatial points. One signal is delayed and multiplied by the spatially separated and undelayed second signal. This half-correlator is combined with another mirror symmetric version. The difference between the two half-correlators forms a motion opponent EMD.

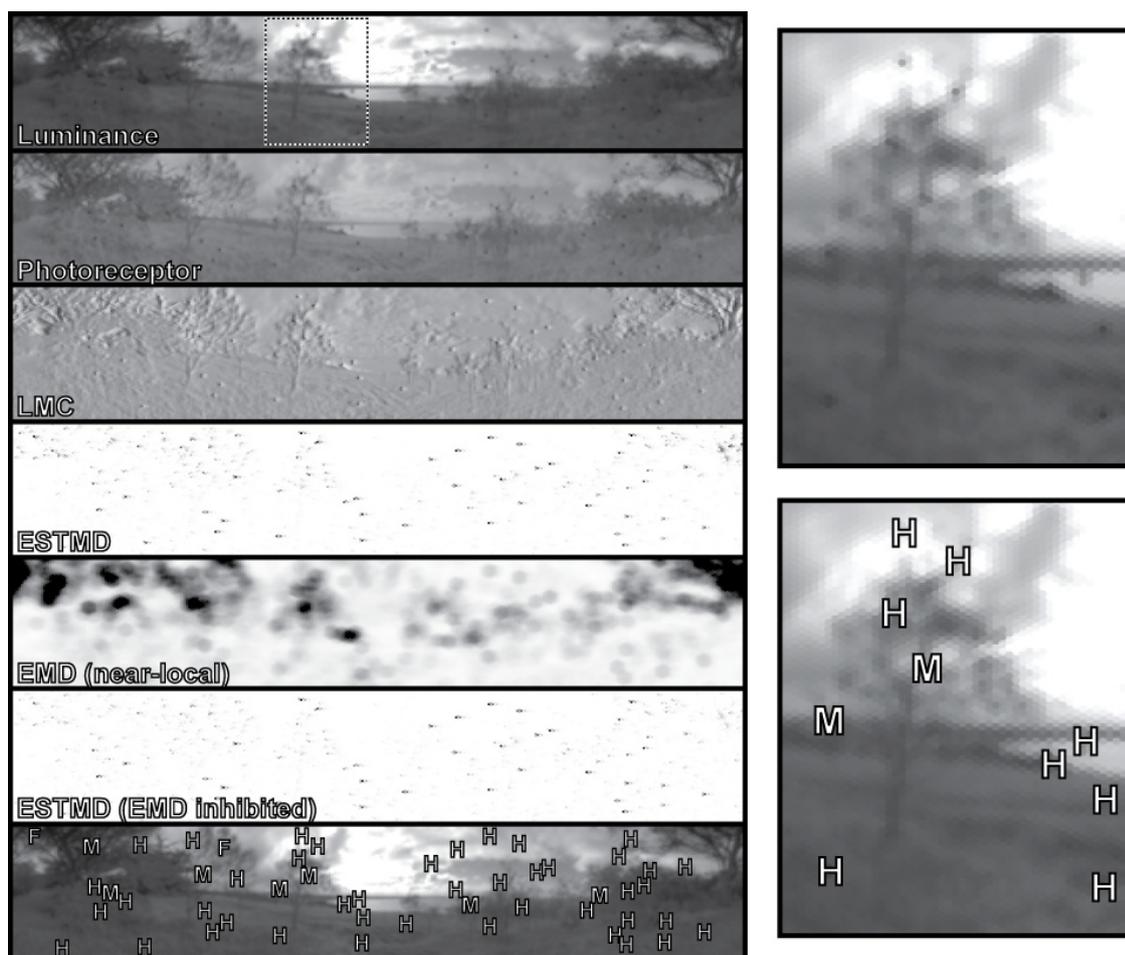


Figure 5-2. Output of the elaborated ESTMD model. Various stages of processing (snapshot in time, panorama velocity of $90^\circ/\text{s}$ with embedded $1.6^\circ \times 1.6^\circ$ targets). ‘Luminance’ shows single green channel (spectral sensitivity), hex sampling (facets) and spatial blurring (compound eye optics). ‘Photoreceptors’ allow for feature discrimination to be viable within a large range of luminance conditions by the use of adaptive filtering and dynamic feedback mechanisms. Large Monopolar Cells (LMCs) extract edge-like features in both space and time with adaptive high-pass filtering. Elementary Small Target Motion Detectors (ESTMDs) extract target information from the LMC input signal by; fast independent polarity temporal adaptation, center-surround antagonism of ‘off’ and ‘on’ channels and a temporal nonlinear facilitation between these channels. The LMC’s also serve as inputs to Elementary Motion Detectors (EMD) and the ‘EMD (near-local)’ is the spatially averaged combination of 37 hex-arranged EMDs (6° across). The ‘ESTMD (EMD inhibited)’ is divisively inhibited by the near-local motion detection which has the effect of sacrificing target detection for reduction of false positives in cluttered, motion rich regions. We determined an optimal strength of this inhibition by varying this parameter in model simulations and selecting the one with best target discrimination. At an optimal threshold we superimpose the ESTMD (EMD inhibited) outputs on the original luminance image [H: hits or true positives, M: misses or false negatives, F: false positives]. This set of images are relevant for display purposes only as they are 8-bit, gamma corrected, representations of the actual model data set.

5.2.3.7 Near-local EMD inhibition

Traditional figure-ground discrimination in insect vision models rely on opposing a local EMD unit with the motion of the surround (wide-field integration of EMD units). This computes a form of relative motion between target and background and there is physiological evidence that some types of figure-ground discrimination rely on this type of processing (Egelhaaf, 1985c, Warzecha et al., 1993). The interesting aspect of the STMD neuron is that it responds robustly, independent of the background motion, even when the target and background velocity are the same. Although the background has some effect on the ESTMD model (establishing adaptive state of the channels, center and surround), this is a limited effect, in contrast to relative motion computations. However, in the model elaborations, we move away from the biomimetic, to examine potential interactions between known physiological elements that would, in an engineering sense, aid in the target discrimination task. Rather than derive model design from the underlying physiology, we are constructing theoretical possibilities to search for in the biological system.

From the output of the basic ESTMD modeling, it is clear that many false positives arise in areas of textured, motion rich surrounds. A method of increasing overall target discrimination is via contrasting our matched target filter output against local EMD-derived motion, i.e. we suppress ESTMD output, with near-local EMD inhibition (across a 6° region). The philosophy behind this approach is that it is as important to suppress false positives as it is to extract target information. In a general sense we are opposing ‘target motion’ against any other form of motion (Figure 2).

To derive an ‘optimal’ amount of near-local EMD inhibition we require quantification of target discrimination. This is determined via analysis of Receiver Operating Characteristic (ROC) curves and calculation of the Area under the ROC (AUROC) (Hanley and McNeil, 1982). From the model outputs at varying stages (luminance, photoreceptor, LMC, ESTMD) we find probability distribution functions of background values and of target location values and use these to construct ROC curves. To ensure a physiologically plausible range of false positives, we equate the 100% false positive fraction as equal to the total number of possible targets, i.e. the AUROC is integrated up to 50 false positives. We ran the simulation with varying strengths of near-local EMD

inhibition and determined the amount that provided the best AUROC results (target discrimination).

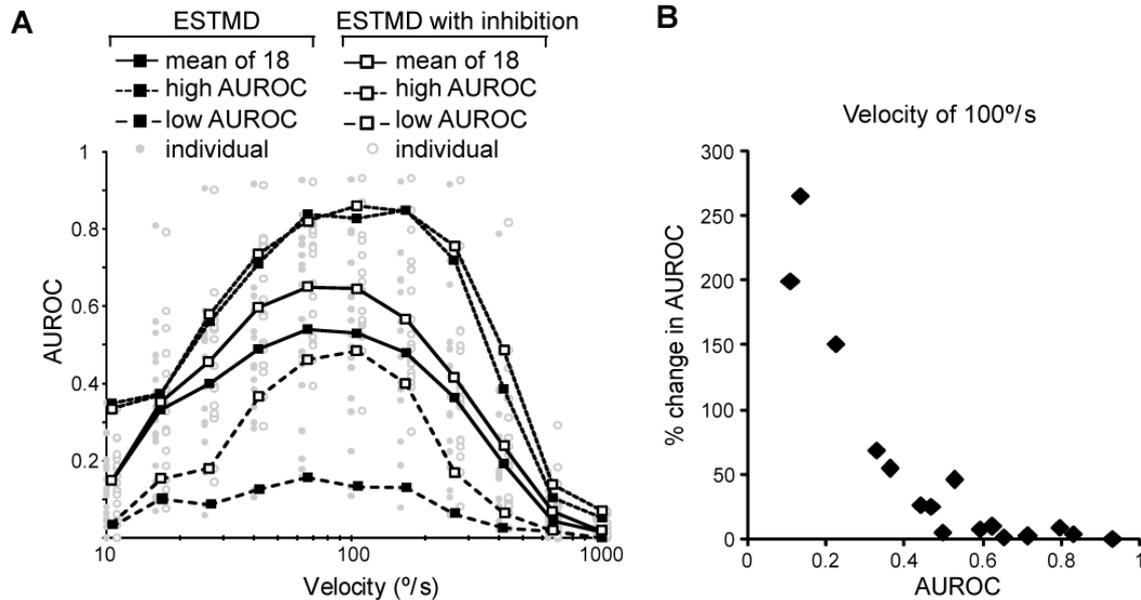


Figure 5-3. ESTMD with inhibition. (a) The Area Under the Receiver Operating Characteristic (AUROC) is a measure of target discrimination and is shown across the operating velocity range for a target size of $1.6^\circ \times 1.6^\circ$. Each natural scene AUROC is shown with the normal ESTMD model (dark circles) and the elaborated version which includes near-local EMD inhibition (open circles, slightly offset for clarity). An image with the second highest mean AUROC values is tracked to represent a ‘high AUROC’ image. Similarly, for the ‘low AUROC’ case (image with second lowest average AUROC values tracked across velocities). There is little difference between the two model versions with respect to the images with high AUROC values (small dashed lines). However, there is a large improvement in target discrimination in the more challenging natural scenes (large dashed lines). For a near optimal example, at $100^\circ/\text{s}$ there are initially 8 of 16 images below an AUROC of 0.5, whilst after near-local EMD inhibition there is only 2 of 16 below 0.5. (b) At $100^\circ/\text{s}$ we plot the percentage increase in AUROC value due to the model elaboration for the 18 images. Images with an initial low AUROC value reveal the largest improvement in target discrimination.

Figure 3 shows the simulation AUROC values across all 18 images, and across the relevant velocity range, with the normal ESTMD model (closed circles) and with this first model elaboration – near-local EMD inhibition (open circles). Although the upper range of AUROCs indicates little improvement (small dashed line), it is in the natural scenes with low AUROC values (large dashed line), i.e. the most challenging natural scenes, where the model elaboration is most valuable. Hence the near-local inhibition model does no worse under conditions where target detection is easy, but does significantly better under more difficult conditions.

5.2.3.8 Ego-motion scaling

Due to the correlation-type nature of the ESTMD model, as the velocity of the panorama (with embedded targets) increases, the entire distribution of output values also increases. At any given velocity, a threshold can be set that provides an ‘optimal’ value for discrimination (Figure 4A), however, if plotted against velocity these optimal thresholds vary across several orders of magnitude (with the underlying shifting distributions). We can account for the shift in these values by determining a reliable estimation of ego-motion velocity and scaling the ESTMD output accordingly.

Figure 4B shows a section of the velocity tuning curve that is derived from the wide-field integration of EMD units and provides a reliable estimation of the panorama velocity (Brinkworth and O'Carroll, 2007, Shoemaker et al., 2005). Two velocity-tuning curves, one representing optimal thresholds (ESTMD), and the other the model-derived estimation of panorama velocity are valid across a physiologically relevant range (though ambiguous beyond this range). The optimal threshold is relevant *even if the targets are not embedded in the panorama scene* as it sets the level of false positives to be contrasted against, rather than establishing a required level for target responsiveness. As both correlator types (ESTMD and EMD) are multiplicative, we took the logarithm of both and found the linear regression equation (Figure 4C). We scaled the ESTMD outputs with a divisive inhibition (via the transform) of a model-derived wide-field velocity estimate.

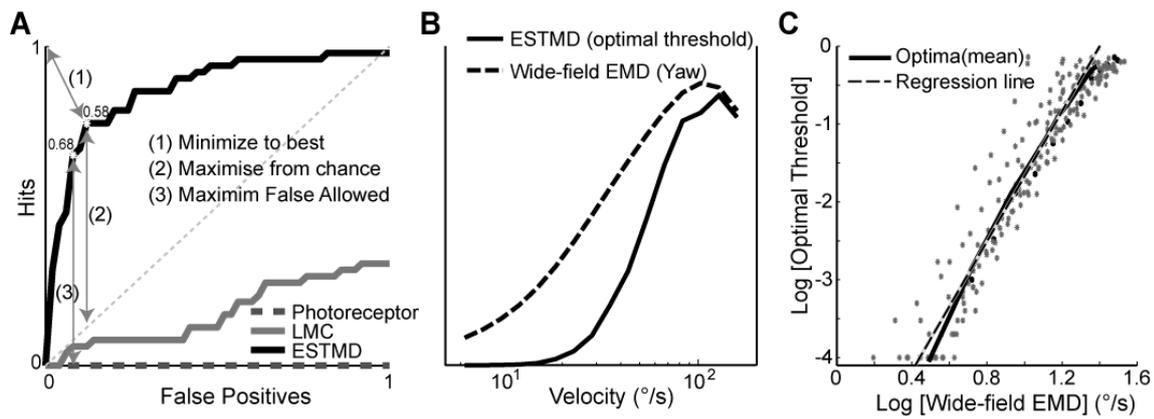


Figure 5-4. Ego-motion rescaling technique. (A) From the Receiver Operating Characteristic (ROC) curve we can determine an optimum threshold value by choosing either: 1) the threshold value that minimizes the distance to ‘best’, i.e. upper left corner with 100% hits, 0% false positives; 2) Maximizes the distance from the equal probability line (with probability of 1 defined as 50 false positives, equivalent to the total number of possible hits); 3) An arbitrary number of false positives (a physiologically plausible scenario). All three values are determined from model simulations and found to be consistent, with option 1 chosen for further analysis and design (B) The mean optimal threshold (minimized to best), across all 18 images at each velocity shows a typical velocity tuning curve (solid line), i.e. the changing speed of the panorama (with embedded targets) increases or decreases responses of both false positives and hits in the same manner, and the optimal value shifts along this curve. A typical panorama velocity (ego-motion) response (e.g. dashed line) produces another velocity-tuning curve. (C) Taking logarithms of both velocity response and the respective optimal threshold values reveals proportionality (due to power relationships inherent in the multiplicative nonlinearities of the two correlation type mechanisms). As ego-motion increases so does the correlation energy inherent in the ESTMD model, increasing the strength of the false positives (and *in the embedded scenario*, increasing strength of targets). Therefore, with a reliable estimation of ego-motion (which is also determined from the model) we can directly scale our ESTMD output to maintain a *static optimal threshold*. With the expected strength of false positives and optimal threshold established, we detect targets exceeding this value irrespective of whether they are embedded or moving relative to the background.

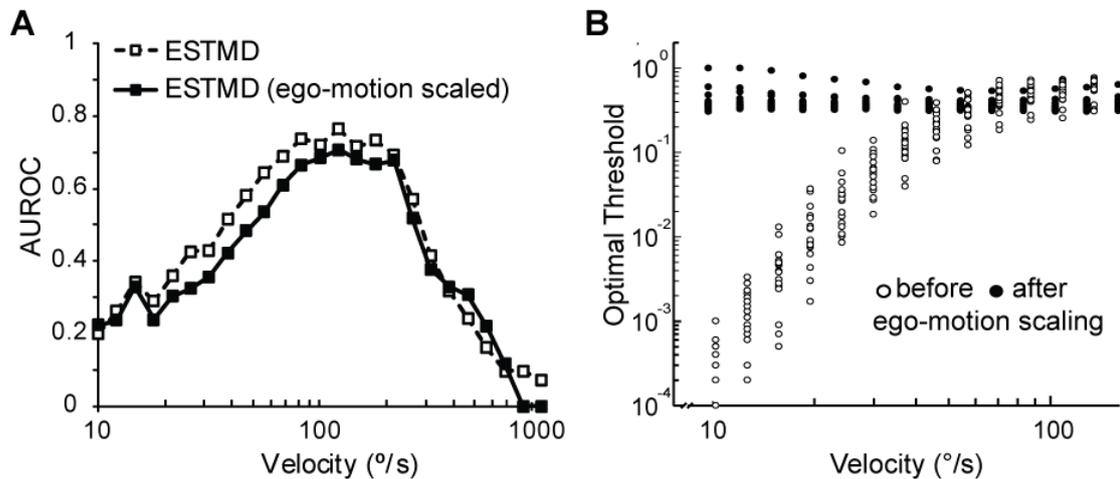


Figure 5-5. Ego-motion rescaled ESTMD. (A) The ability to discriminate targets (AUROC) is similar between the ESTMD model and following ego-motion scaling (same image as shown in Fig 2, 1.6°x1.6° targets). At any single panorama velocity, the ESTMD output is divisively inhibited with a scalar value (therefore not changing the AUROC). (B) Following ego-motion scaling, the optimal threshold values of the 18 natural scenes, remain relatively constant (closed circles), whilst in the normal ESTMD model the optimal threshold varied over many orders of magnitude, dependent on panorama velocity (open circles).

To verify that this had no effect on the ability to discriminate targets, we show the similar AUROC values for a given image under both normal conditions, and following ego-motion rescaling (Figure 5A). We expect the same AUROC values as for any particular panorama velocity we are simply scaling the distributions by a scalar value (small differences are due to methodologies and simulation approximations). Figure 5B reveals that before ego-motion scaling the optimal target discrimination threshold is distributed over many orders of magnitude and that following the operation they are now relatively constant.

Therefore, we can have a single (and optimal) value for target discrimination that is applicable under varying amounts of ego-motion. In effect, the velocity dependent ESTMD excitability of the overall system is being normalized, and this allows the target motion discriminator to work within a static range (and with a static optimal threshold).

Figure 6 shows the transform in finer detail, with cumulative distribution functions (for the varying velocities) that shift after ego-motion scaling. As a suitable approximation, we chose this simple transform, however, a higher-order curve fit could establish a more accurate ego-motion scaling function (Figure 6E; lines could overlay).

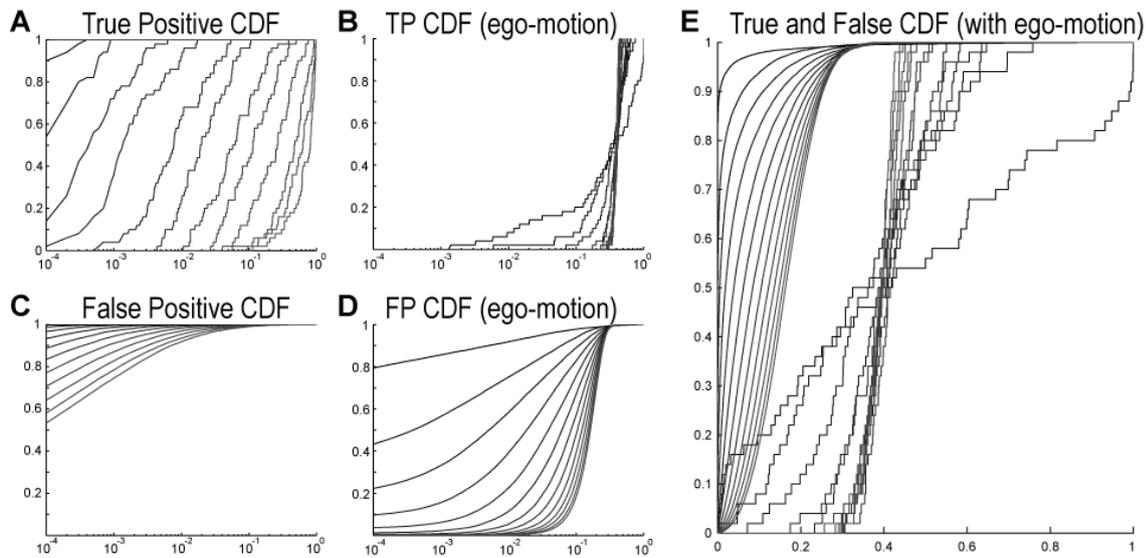


Figure 5-6. Cumulative Distribution Functions. (CDF) at varying velocities (lines shift to the right as velocity increases). (A) CDF of target location values in the ESTMD model and (B) following ego-motion rescaling (note the logarithmic scales). (C) We show output values at all non-target locations (to indicate false positives) for the normal model and (D) under ego-motion scaling. In (E) we see the resultant effect of ego-motion scaling (from B and D, on a linear axis) which allows for both a static optimal threshold and maintains separation of the False Positive and True Positive CDFs, i.e. target discrimination (as also evidenced in Fig. 5A. AUROC values)

5.2.4 Conclusion

We find that near-local EMD inhibition increases target discrimination by suppressing false positives in highly cluttered and motion-rich regions. This design is similar to previously published models (Egelhaaf, 1985c, Higgins and Pant, 2004) where local EMD units are inhibited from wide-field integration of EMDs. However, in our modeling it is inhibition of target motion (ESTMD output) by the integration of surrounding EMD units. We have established an optimal amount for this suppression by varying the strength of the inhibitory influence and then calculating the respective AUROC values until target discriminability is maximized. This optimal value equates the benefits and costs of hits and false positives, whilst in a biological system these aspects are most probably unbalanced. As prey, this weighting should be determined from the metabolic costs of avoiding moving objects versus the cost of survival. As predator, the metabolic cost of target pursuit versus the required territorial and mating chase behaviors.

As we increase the spatial scale of integrated EMD inhibition, we find that the wide-field (velocity estimator) inhibition serves to scale the ESTMD output to take into account overall ego-motion excitability. This allows for the implementation of a static optimal threshold for target discrimination under varying values of ego-motion.

In these simulations we have dealt with the target embedded in the panorama and therefore without relative motion cues. Previously, in the non-elaborated model we have shown that target responses are robust, independent of relative motion between target and background (Wiederman et al., 2008). That is, the target speed determines the strength of ESTMD response, even in the presence of complex moving backgrounds.

In the case of the elaborated model, the inclusion of motion antagonistic components will have important modulation effects on ESTMD output, whilst still allowing for the robustness of response without relative motion cues (as observed in physiology). We can test for these modulation effects in future electrophysiological experiments. If target speed across a detector is low, in an environment of high ego-motion, then we would not expect a strong neural response. Although intuitively this might be considered a model design constraint, in some scenarios this information from ‘low’ ESTMD responses could be a plausible neural approach to target pursuit. As long as a position dependent response had indicated the presence of a target then increased ESTMD responses would indicate increased motion of the target away from the center of pursuit. Further research will investigate such visual stimulus conditions, more akin to those found in insect visual ecology (with differing target speeds in the presence of translational and rotational ego-motion components).

From the engineering perspective, we have elaborated the model to examine potential implementations of these bio-inspired processes in engineering applications. From the physiological perspective, the proposed elaborations allow new stimuli to be used in electrophysiological experiments in order to investigate whether such elaborative processes are found within the neural systems.

5.2.5 Acknowledgements

This research was funded by the US Air Force Office of Scientific Research (FA 9550–04–1–0294) and by the Australian Research Council in the form of a Linkage Grant (LP0667744) and Post Doctoral Research Fellowship (RSAB).

CHAPTER 6:

TESTING THE ESTMD MODEL WITH EXPERIMENTS ON STMDs

6.1 CONTEXT

This chapter is a manuscript in preparation for publication. In final form it is possible that this will be split into two manuscripts. At present it lacks the extensive background information that will be included in the final manuscript. This has already been provided in earlier chapters so I have omitted it here to avoid repetition.

6.2 COMPARISON OF THE ESTMD MODEL TO THE DRAGONFLY CSTMD RESPONSES

6.2.1 Introduction

The central aim of this thesis was to develop a robust and simple model for target discrimination in visual clutter based on specific tuning for features that are uncommon in typical natural scenes. The model presented in the earlier chapters provides a close fit to the basic tuning of typical fly or dragonfly STMDs in terms of their size tuning, velocity tuning and possibly even their sensitivity to contrast (i.e. preference for negative contrast). However, to verify whether this model captures processing typical of that in biological STMDs it is important to test its predictions in the presence of visual clutter against the quantitative responses of biological STMDs.

I therefore performed intracellular recordings from CSTMD (centrifugal) neurons within the optic lobe of the dragonfly (*Hemicordulia tau*). These neurons have been shown to exhibit classical STMD type responses such as size selectivity, contrast sensitivity, and velocity tuning (O'Carroll, 1993, Geurten et al., 2007). While many STMDs in insects are small neurons and therefore difficult to record, the CSTMD neuron is large and easily recorded for prolonged periods. It therefore provides an ideal basis for quantitative comparisons between model predictions and biological responses.

The aim of the experiment was to examine neural responses to natural scenes, in particular to an embedded target and responses to 'false positive' features, as predicted by our ESTMD modelling. In addition, by subtle digital manipulation of the images, I removed some of the model predicted 'false positives' from the visual stimulus thus generating testable predictions for variation in the CSTMD responses.

6.2.2 Methods

6.2.2.1 Electrophysiology

Dragonflies (*Hemicordulia tau*) were caught in the wild and then immobilized with wax. A small hole was cut in the back of the head and the cuticle removed. Air sacs and other tissue were removed to provide clear access to the left lobula. The perineural sheath was left intact. The dragonfly was positioned approximately 15cm away from a

200 Hz CRT monitor, mean luminance of $100 \text{ cd}\cdot\text{m}^{-2}$. The visual stimuli were programmed in Python, using the VisionEgg stimulus software (www.visionegg.org).

Micropipettes were pulled from 1 mm (O.D.) thick walled alumina-silicate glass capillaries (SM100F-10, Harvard Apparatus Ltd.), on a Sutter Instruments P-97 puller, and filled with 2 M KCl or 2 M potassium acetate. Electrode resistances were typically 120-150 M Ω .

The data were sampled at 5 kHz during acquisition, using a National Instruments 16-bit ADC. Data analysis was performed offline with Matlab.

6.2.2.2 CSTMD Identification

Neurons were recorded intracellularly in the left lobula of the brain. To identify the neuron as CSTMD, I mapped the receptive field (Figure 6-2). This was done by drifting a small black target ($0.8^\circ \times 0.8^\circ$) at $50^\circ/\text{s}$ across the white screen in a series of 21 vertical and 21 horizontal scans (in both directions). I then determined size selectivity of the neuron by scanning across the monitor a series of variable height bars (at 0.8° width). The CSTMD gives no response to wide-field stimuli, such as sinusoidal gratings.

6.2.2.3 Number of Trials

Although I obtained partial recordings from a large number of animals, data presented here were obtained from two dragonflies in which the recordings were of long duration (total duration of 16 hours) and which were stable enough for the lengthy stimulus protocols. During prolonged recordings a receptive field scan (see 6.2.2.2), performed every two hours, allowed for testing of the health and viability of the cell. These scans revealed that the functionality of the entire CSTMD pathway was intact and healthy. Over this time, spontaneous spike rate varied, however, target responsiveness remained similar (maximum rate of $\sim 250\text{-}300$ spikes/s).

6.2.2.4 Image preparation

Two of our panorama natural images were selected (4.2.3.2), one ‘natural’ and the other ‘urban’ (Figures 6.10-6.11). The images were cropped to a limited vertical extent (10°) and the top and bottom edges were altered to gradually fade to the background grey.

I ran these two images (without targets) through the ESTMD model (as previously described) so as to determine the location of the false positives. By using the Adobe Photoshop™ ‘healing hand’ tool, I was able to ‘remove’ these features in the image.

Via this digital manipulation of the natural images, I constructed three different versions of the same scene; an unaltered image, the image with one false positive removed and the image with an embedded target.

6.2.3 Results

6.2.3.1 Spiking Receptive Field Mapping

I mapped the receptive field, initially to identify the neuron as CSTMD. As I was to present natural scenes (with limited vertical extent) to the dragonfly, it was important to determine the region of the receptive field that contained the highest excitatory responses (the ‘hotspot’) and to examine the spatial homogeneity of this field. The presentation of the natural images could then be centred in this region.

In Figure 6-1 we see a single scan of the receptive field mapping to the $0.8^\circ \times 0.8^\circ$ size target. A clear inhibition of the spike rate is evident as the target drifts across the left side of the dragonfly midline (ipsilateral side), i.e. in the visual field of the other eye. As the target crosses the midline and enters the contralateral side, the spiking rate increases.

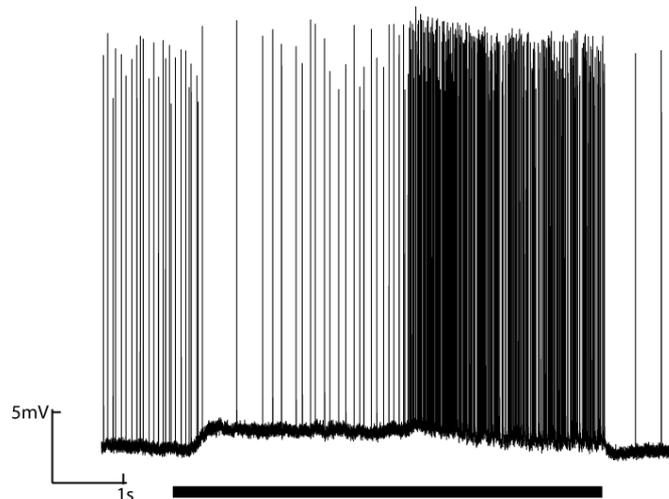


Figure 6-1. Single Scan through Receptive Field. This response is to the $0.8^\circ \times 0.8^\circ$ target used in the receptive field mapping (Figure 6-2). The stimulus duration is indicated by the black bar. Soon after the target enters the monitor screen the spiking is inhibited whilst the membrane potential depolarises. As the target crosses the midline, the hotspot is reached (in the contralateral eye) and spiking rate increases.

Figure 6-2 shows the averaged spiking response over multiple trials. Note the excitatory receptive field runs along the midline of the dragonfly, between the two eyes (at 0° azimuth). The CSTMD neuron is weakly direction selective, i.e. responding in both directions, but preferring progressive target motion (upwards and to the right). Therefore, presentation of the natural scenes should be animated at a chosen velocity from left to right on the CRT monitor, centred on the hotspot.

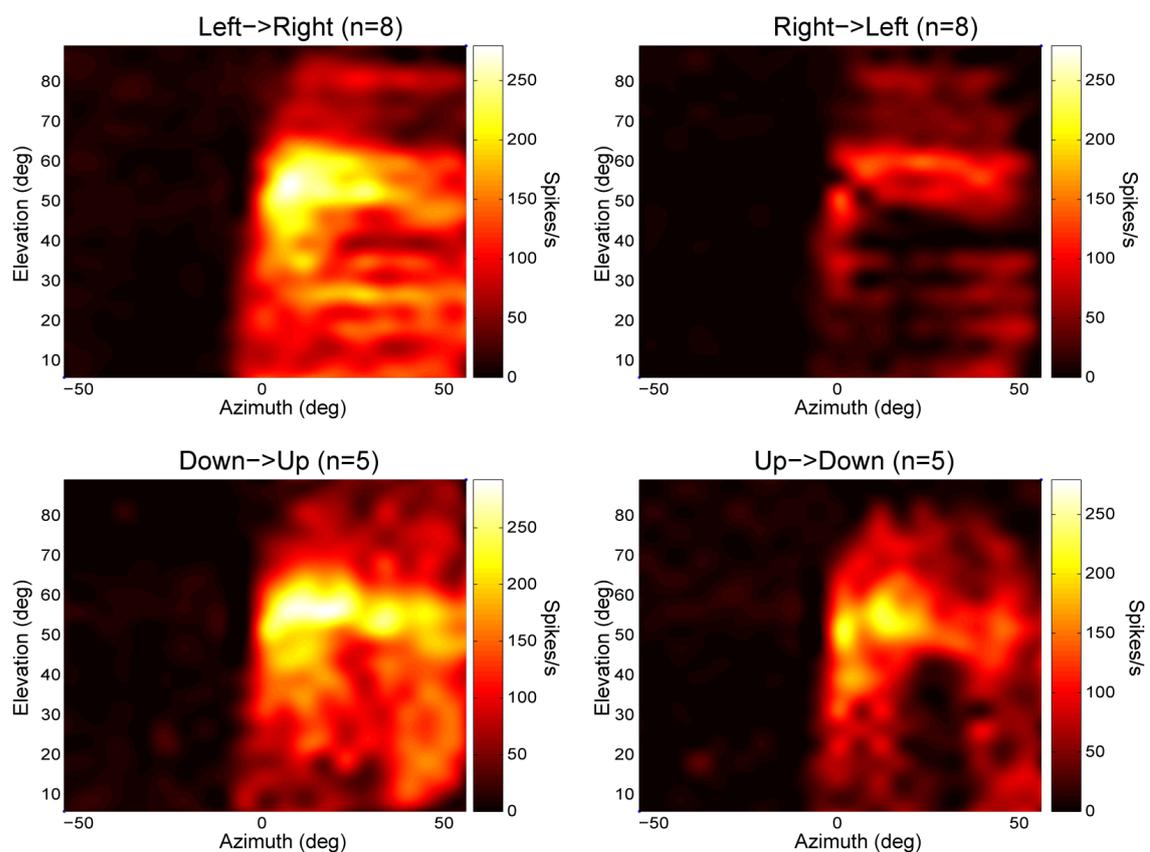


Figure 6-2. Spiking Receptive Field Scan. A small black target (0.8°x0.8°) is drifted across the white monitor screen and the strength of the spiking response is graphed. The response begins spiking as the target crosses the midline of the dragonfly eyes. The response is stronger to progressive target motion (up and to the right).

The response to a near optimal target size reveals a high spiking rate over a large region of space. However, this ‘field of effect’ will vary dependent on the nature of the stimulus. As I will examine the responses of CSTMD to the natural scenes (each containing many variations of stimulus features), it is prudent to examine how the receptive field region changes in response to a different size feature.

I repeated a receptive field mapping with a single pixel stimulus ($0.16^\circ \times 0.16^\circ$). The interommatidial angle for *Hemicordulia tau* is 0.56° (Horridge, 1978), therefore this sub-pixel target is below the spatial resolution of the eye and thus is effectively a single ommatidium stimulus of low contrast. The neuron responded strongly in a small part of the visual field, corresponding to the hotspot in the previous mapping. This single trial experiment shows high contrast sensitivity for the dragonfly CSTMD.

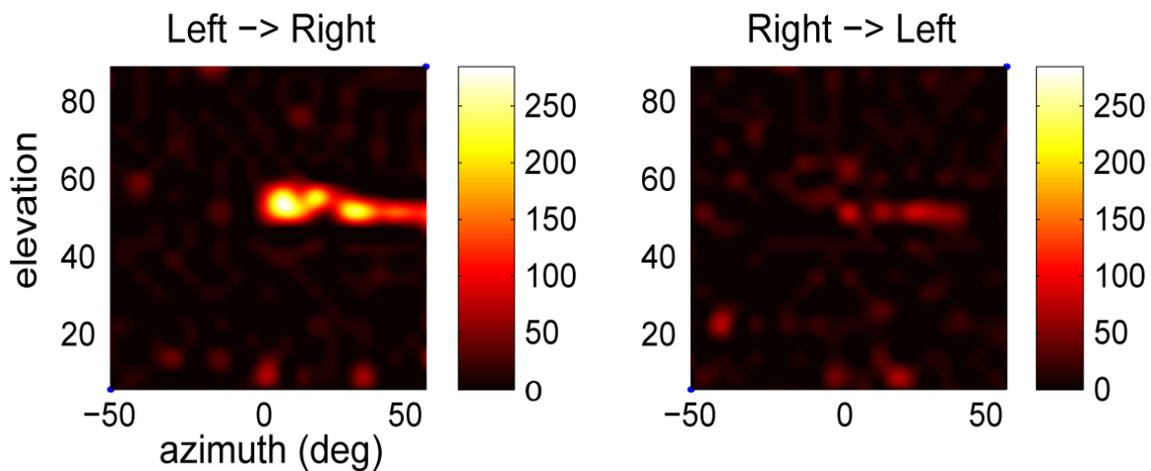


Figure 6-3. High Contrast Sensitivity. Spiking response (spikes/s) to a drifted target of only $0.16^\circ \times 0.16^\circ$ ($N=1$). As the target is smaller than the interommatidial angle, a target of this size is effectively one of low contrast. This reduced contrast target reveals the strongest ‘hotspot’ of the CSTMD receptive field.

6.2.3.2 Graded Receptive Field Mapping

A further interesting result is seen in the graded generator potential of CSTMD during the receptive field mapping. An example scan is shown in Figure 6-1. Soon after the target appears on screen (see the stimulus duration bar), a depolarisation is observed. This occurs whilst the target is on the ipsilateral side and the spike rate is inhibited below the spontaneous rate.

By using a spike removal algorithm, I determined the changes in this underlying potential and plotted a ‘graded’ receptive field map (Figure 6-4). These graded shifts were only evident during a limited period of the recording time, so the strength of their effect has been averaged out over the entire experimental duration. However, qualitatively this result does indicate that the changing voltage is present across a large region of the visual field (both the ipsilateral and contralateral side). Surprisingly, this graded response to the target is a depolarising shift that coincides with either the spiking

inhibition (ipsilateral) or the spiking excitation (contralateral). During the excitatory stage the generator potential is driven back towards the baseline potential.

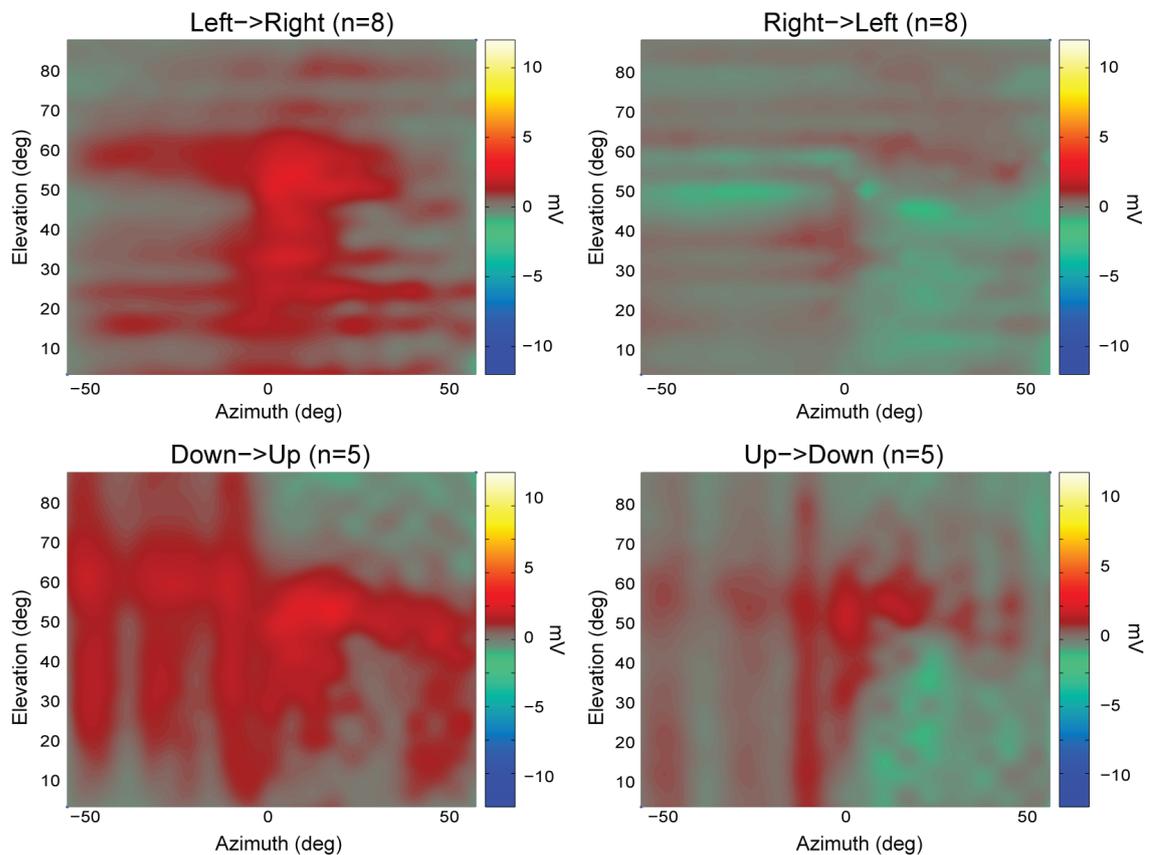


Figure 6-4. Graded Receptive Field Scan. The graded response (underlying change in generator potential) to a $0.8^\circ \times 0.8^\circ$ target. The amount of graded response varied throughout the experiment, however, when present, involved depolarisation at the same time as inhibition (ipsilateral) and excitation (contralateral) of the spiking response .

6.2.3.3 Contrast Polarity Selectivity

An inherent property of my ESTMD model is that it is selective for polarity of contrast i.e. responsive to dark targets and not to an equivalent contrast light target. In addition, as the natural images to be presented to the CSTMD will have both light and dark features within the one scene, then determining any contrast polarity selectivity of the CSTMD neuron will be invaluable in interpretation of those results.

I investigated contrast polarity selectivity within the CSTMD, by presenting targets on a mean background (RGB value 0.5), and then varying the target RGB values, so as to present both light and dark contrasting targets (Figure 6-5).

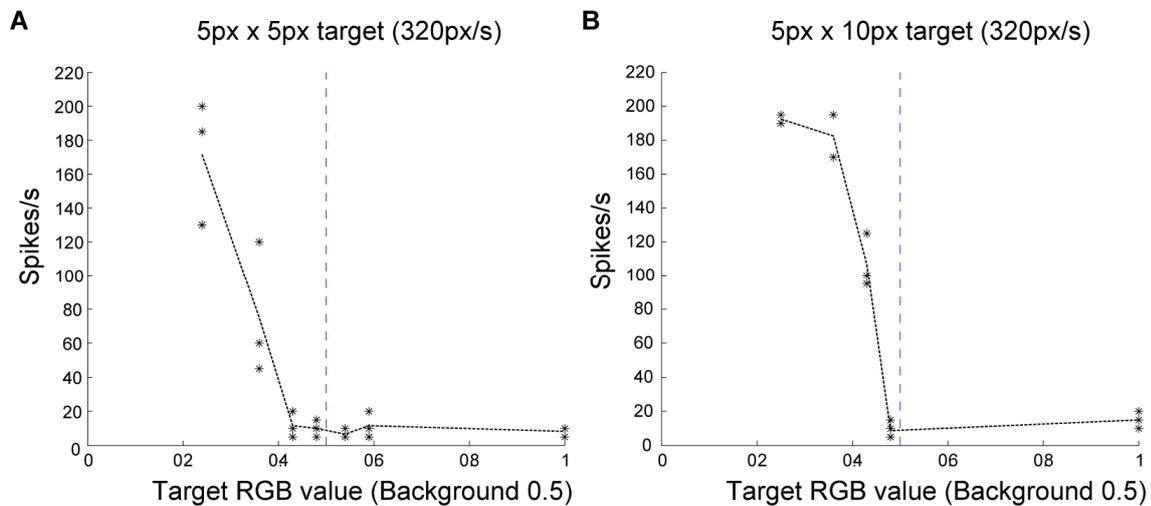


Figure 6-5. Dark Target Selectivity of CSTMD. A) The neuron response to varying target contrasts ($0.8^{\circ} \times 0.8^{\circ}$). The response to equivalent contrasts (RGB 0.24 - dark and 1 - light) reveals strong polarity selectivity. The dashed midline represents background RGB value. As target RGB value decreases below 0.5, target dark contrast increases. There is limited response to lighter targets than background. B) Repeated experiment with increased target height (1.6°). Three repeats at each RGB value, a line joins the means.

An RGB value of 1 (white target) on a 0.5 background is the equivalent Michelson contrast (Michelson, 1927) as a value of 0.24 (dark grey) on the same background. Given a more appropriate contrast measure for targets on background (Weber contrast) and these RGB values result in contrasts of 0.52 (dark target) and 1 (white target). The CSTMD neuron is even unresponsive to a white target of double the contrast.

I further verified this contrast polarity selectivity using both a dark target and a light target and in each case varying the target height (Figure 6-6). These results also show that regardless of the height of the white target, the CSTMD neuron will not respond to lighter targets.

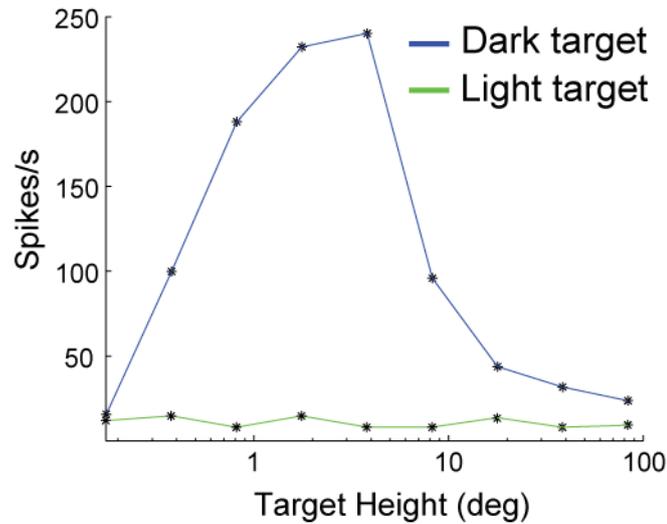


Figure 6-6. Target Height Tuning (Dark and Light Polarity). Target height sensitivity is shown for a dark target (RGB of 0.24) on a mean background (0.5), peaking at $\sim 2.5^\circ$. Also on the mean background, an equivalent contrast light target (RGB of 1), with varying height is shown. Again, dark target selectivity is revealed.

6.2.3.4 On and Off Channels in the Insect Visual System

Previous intracellular recordings from the medulla of the blowfly (DeVoe and Ockleford, 1976) showed neural responses to a step increase in light intensity. To this stimulus the authors observed both neurons that increased in spike rate and neurons that decreased in spike rate. This indicates a potential separation of channels within the insect vision system.

Figure 6-7 provides further evidence for this type of processing in an intracellular recording from a neuron in the medulla of the blowfly. The neuron has a decreased response to light increments, and then a sustained, increased spiking response to a step decrease in luminance level. This type of processing could underlie the rectification required for the previously described contrast polarity selectivity observed in the CSTMD.

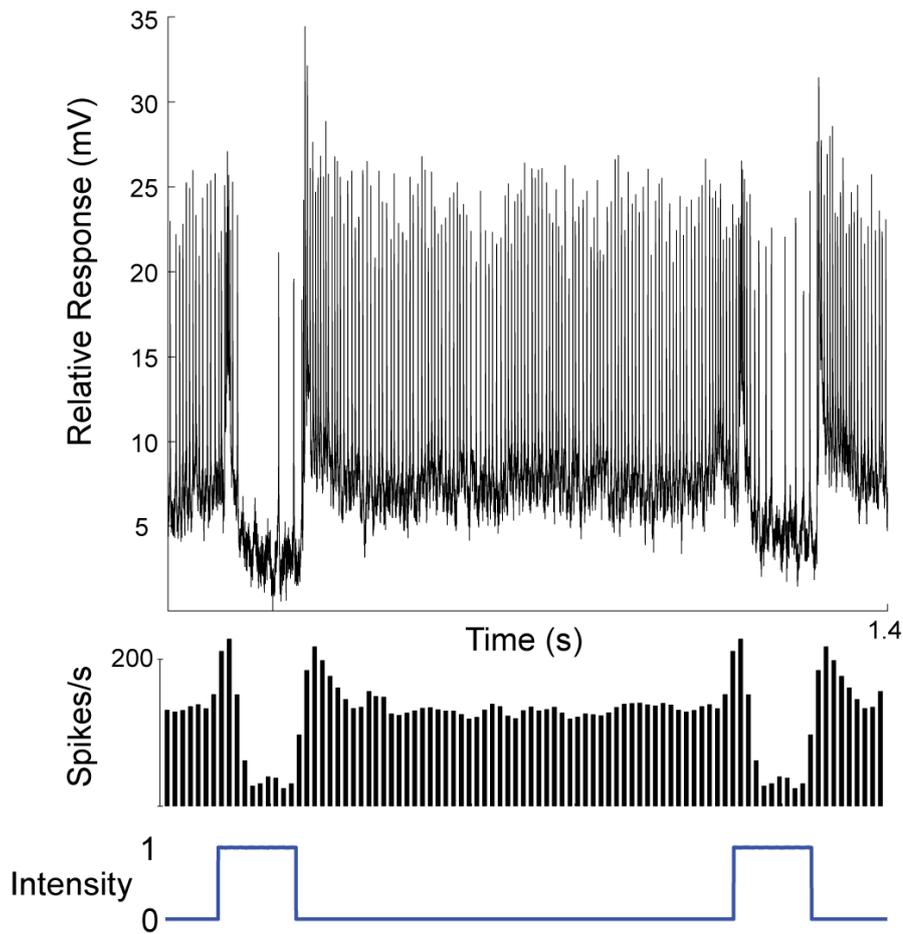


Figure 6-7. An Off channel in the medulla of the blowfly, *Calliphora*. This intracellular recording shows a decrease in response to a step of increased luminance (from a high power LED). The response to lower intensity is sustained and spiking. Note that the transient changes show a full-wave rectified depolarization of the response. The spike histogram is the average over 10 cycles of the stimulus.

6.2.3.5 Natural Image Experiments

The ESTMD model is tuned to detect a spatiotemporal signature that is quite rare within natural scenes (see 2.2.4.9). However, given a natural image input, the model will generate some false positives. If we include targets within the natural image then, as expected, the model is responsive to those targets.

The question then arises as to whether, given a similar natural image stimulus, the CSTMD neuron responds in the same manner? Is an embedded target within the scene a driver of the neuron and does the CSTMD respond to the same false positives in the same way? Furthermore, if we manipulate the image to remove the model false positives, can we induce the same response in the CSTMD?

In Figure 6-8 we see the ESTMD model responses to a region of the natural image variants. The first panel (A) shows the unaltered image and (B) the model response. In panel (C), a single false positive has been digitally removed from the natural scene and in (D) the model output no longer shows a response in that location. In panel (E) a target has been inserted and in (F) the strongest ESTMD response is to that target.

Therefore, the ESTMD model identified false positives within the natural scene. One of these false positives (within the entire image) was removed via digital manipulation (the ‘leaves’). The model simulation was run again, following the image alteration, to verify that this location no longer contained a false positive. Finally, we checked that the model response to target insertion was strong.

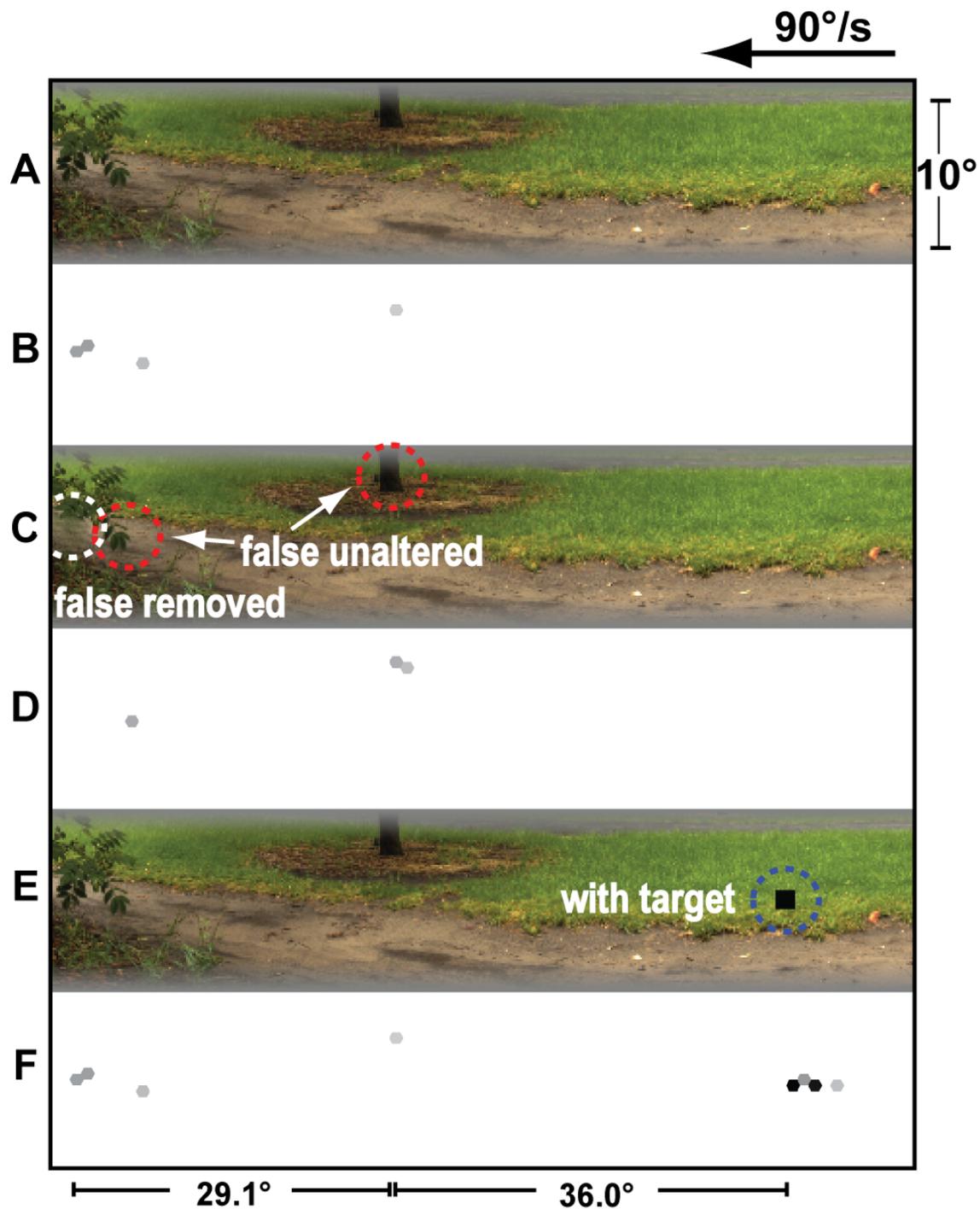


Figure 6-8. ESTMD Model Responses to the Natural Images. An unaltered natural scene (A) is input to the ESTMD model and the output (B) determines the false positives. A single false positive is digitally removed from the scene (C) and the model no longer responds in that location (D). Finally, I embed a target in the scene (E) and this provides the strongest model response (F). The panorama velocity is at $90^\circ/\text{s}$. Some example distances are also shown.

6.2.3.6 A Note on Space and Time

The visual panorama stimulus is to be presented on the monitor and rotated past the dragonfly in the direction that will stimulate the directional CSTMD neuron (left to right). To portray the spiking responses of the neuron to this stimulus (over time), so that it correlates to the visual features within a scene (over space), we are left with no alternative but to either have the time axis presented from right to left (unconventional), or to provide a mirror symmetric representation of space. The later was chosen for the presentation of this data, however, this entails no significant changes to the interpretation. The presentation of all spatial dimensions (Figure 6.8-6.11) has been mirrored and the direction of velocity reversed. It is similar to what the ‘receptive field’ would look like from the CSTMD neuron on the other side of the brain (Figure 6-9).

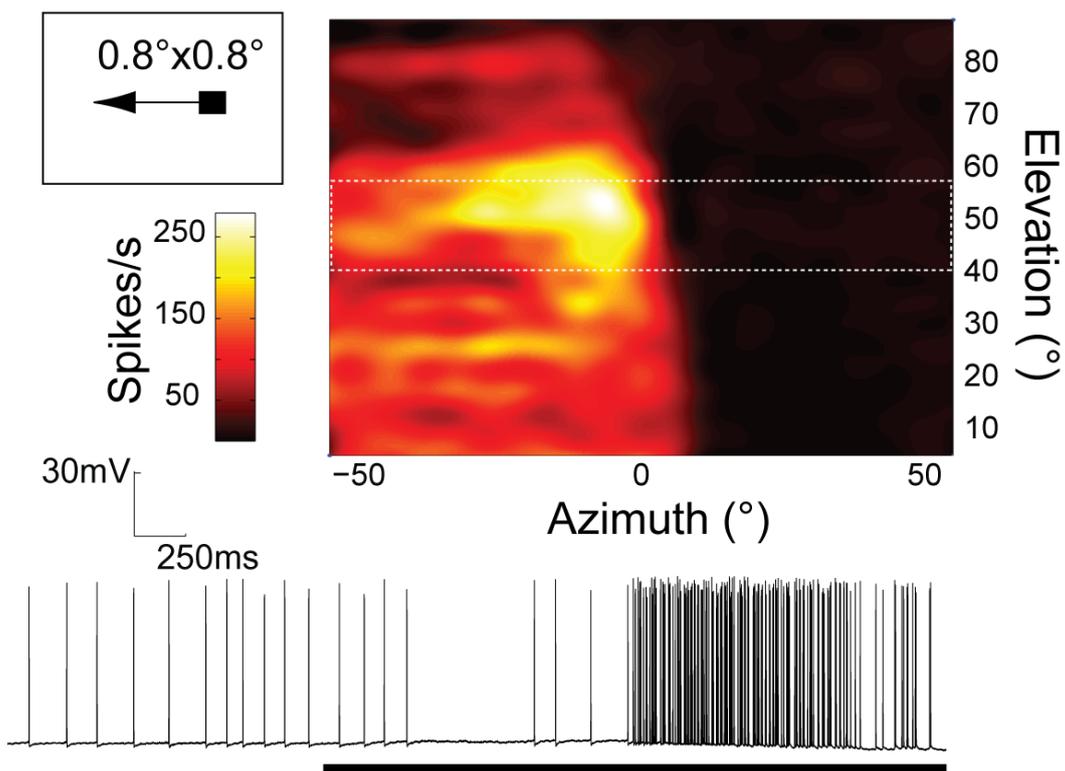


Figure 6-9. Mirrored Receptive Field. The receptive field (mapped by the drifting target) is shown mirrored around the vertical midline. This allows for easier interpretation in our analysis of the neuronal responses over time. The white dashed lines represent the size of the natural scenes that are to be presented to the dragonfly. The neural response in this region is relatively homogenous, though features within the bottom part of the natural images are less likely to drive the neuron.

6.2.3.7 CSTMD Electrophysiological Responses to the Natural Images

To see whether the CSTMD neuron responded in the way predicted from the ESTMD model output, I recorded intracellularly whilst stimulating the neuron with the series of natural images. I displayed these panoramas on the CRT monitor, centred on the hotspot derived from the receptive field mapping.

Rotating the panoramas at 45°/s, in the preferred direction of the CSTMD ensured that we were driving the neuron with a near optimal velocity (Geurten et al., 2007). The ESTMD false positive predictions were simulated at 90°/s with a spatial sampling resolution of the fly, i.e. approximately 1°. With the 0.56° resolution of the dragonfly, the two different velocities (doubled), coinciding with the different spatial sampling (halved) and therefore should be roughly equivalent.

The CSTMD neural responses to the visual stimulus are shown in Figure 6-10 and Figure 6-11. The spiking responses of the neuron are shown in both spike raster plots and spike histograms (mean \pm SD). The time dimensions of the neural responses are aligned with the images so that it is coincident with the 0° azimuth in Figure 6-9 (near the hotspot). At any instant, a \sim 110° region of panorama is shown on the monitor screen. The CSTMD receptive field spatially integrates over this region, as represented in the ‘mirrored’ version in Figure 6-9. Note that the target is embedded in the scene, with no velocity disparity between the target and background.

In image ‘Bush’ (Figure 6-10), we first examine the neural responses to the original scene (1st panel). There are intermittent responses to the ‘leaves’ (the false positive to be removed, black arrow), but also a somewhat sensitive response as the tree trunk (and another feature) drags across the receptive field. In the second panel, the ‘leaves’ have been removed and there is a dramatic reduction in response. This is not an effect of habituation as all of these experiments were pseudo-randomly ordered with unadapting periods between the trials. In some more complex manner, the responses to the tree trunk are also gone (possibly a reduced sensitivity). In the third panel we have some intermittent responses returned to the false positive ‘leaves’ and now a strong response to the embedded target (red arrow). Any lack of neural responses to the tree trunk are now readily explained by the inter-ocular inhibition, as the target drifts through the inhibitory field before and during the passage of the tree (see distances in Figure 6-8).

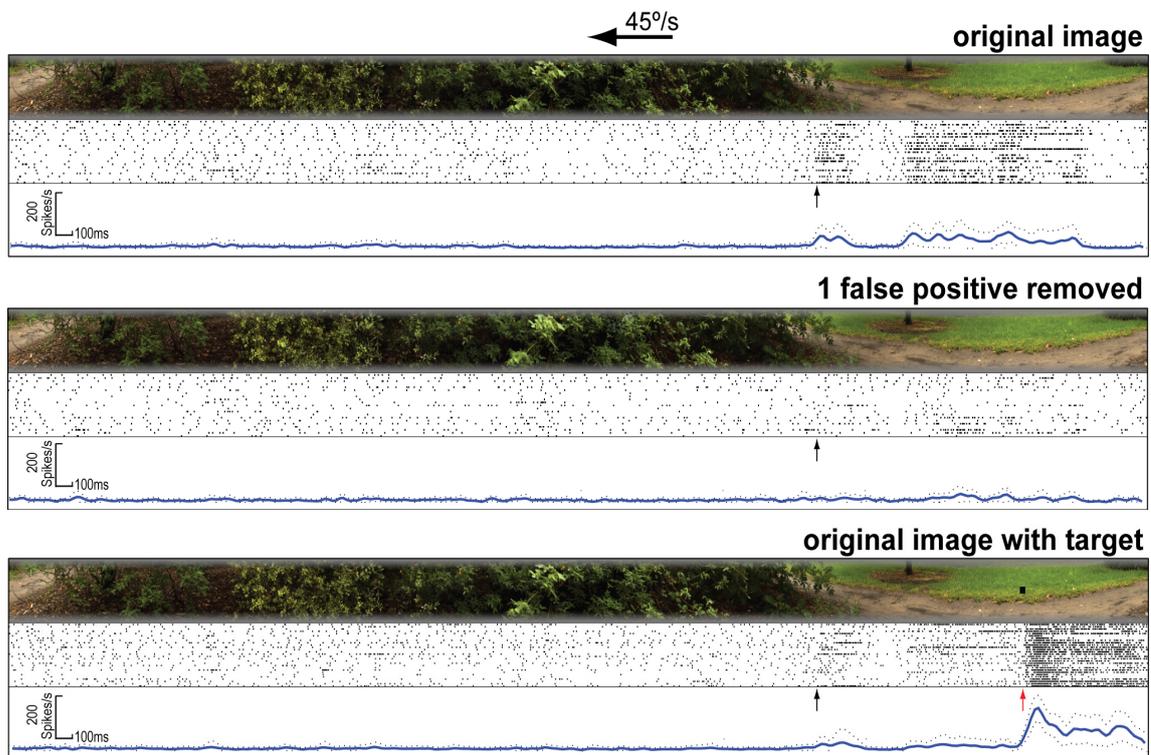


Figure 6-10. CSTMD Neural Responses to the ‘Bush’ Images. Images are rotated past the dragonfly (at 45°/s) and CSTMD neural responses (over time) are aligned with the image spatial dimension (centred at 0°, see Figure 6-9), so that correlation with features is shown. Monitor width displays ~110° of the 360° panorama, at any instant in time. The trials were not ordered (pseudo-randomly distributed between all six images in Figure 6-10 and Figure 6-11) and pauses between the trials took into account adaptation effects. See in text for a description of the responses.

In image ‘Car Park’ (Figure 6-11), we observe similar characteristics, with the same effect occurring by removing the false positive (black arrow). However, the CSTMD is tuned for slightly larger height selectivity (Figure 6-6) than the ESTMD modelling (Figure 2-7) and often intermittently responds to the passing columns (blue arrows).

The response to the target in the third panel (red arrow) is more transient in nature (green arrow) in contrast to the target in image ‘Bush’. This is again explained by the influences of features within the inhibitory field suppressing the target response, an effect that does not occur in Figure 6-10.

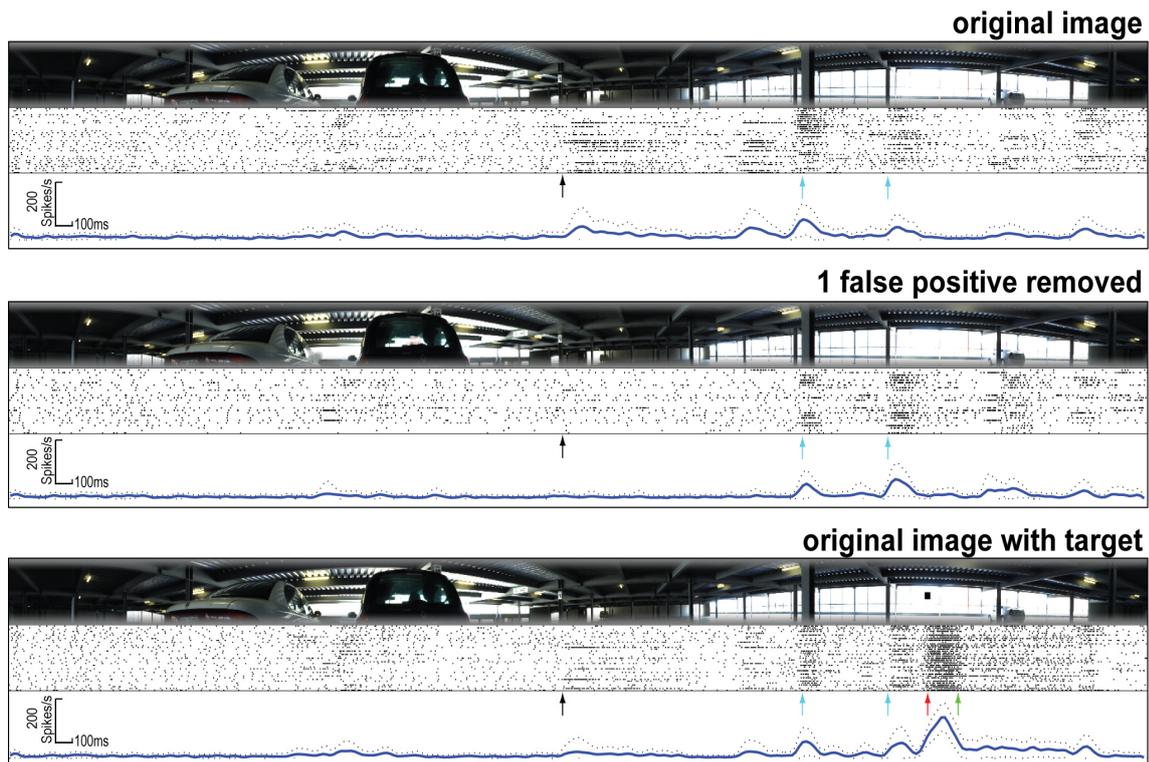


Figure 6-11. CSTMD Responses to the ‘Car Park’ Images. Details as in the previous figure.

6.2.4 Discussion

The results of the CSTMD experiments provide some supporting evidence for an ESTMD-type framework (or equivalent) underlying the STMD mechanism. I showed that the CSTMD is responsive to feature types embedded within a scene, without the presence of any global velocity disparity between the features and background. Locally, motion energy is inhomogeneous and could still form an important cue for feature discrimination (5.2.3.7).

The high contrast sensitivity (Figure 6-3) in response to the single pixel target ($0.16^\circ \times 0.16^\circ$) reveals that a non-optimal stimuli (small size, low contrast target) can generate a powerful response at an earlier stage of the visual pathway. The only difference to non-optimal stimuli in the natural scenes is that there is the presence of some contrasting surround (rather than a white background). This indicates that to produce the target selectivity there must be a powerful inhibitory component. Within the ESTMD model, this is via the strong spatial antagonistic surrounds (2.2.4.3) or in the elaborated model (see 5.2) via a motion-derived near-local antagonism.

Whilst mapping the receptive field, the inter-ocular interactions were immediately evident. This same result was indicated in earlier physiological experimentation (Geurten et al., 2007) and understanding these inhibitory interactions is presently a research focus of an ongoing PhD project in this laboratory (Douglas Bolzon, in preparation). The CSTMD is a higher-order neuron (complexity of visual processing) than what is observed in the STMD of the hoverfly.

Sometimes during intracellular recordings, subtle changes in electrode position can provide deeper insight into the neuronal physiology. This was evident for a while in my experiments, as I was able to observe the changes in the underlying potential. At this stage, the recording of the neuron must have been far from the spike initiation zone, but experiencing a depolarisation (due to the presence of a target on the ipsilateral side). This influence extends too far to be due to any form of binocular overlap (maximum of 10°) (Horridge, 1978)

In Figure 6-12, we see a diagram of CSTMD morphology (reproduced from Geurten et al., 2007). From their experiments, the authors predicted that interactions between a symmetric pair of CSTMD neurons would be the likely cause for this inhibition.

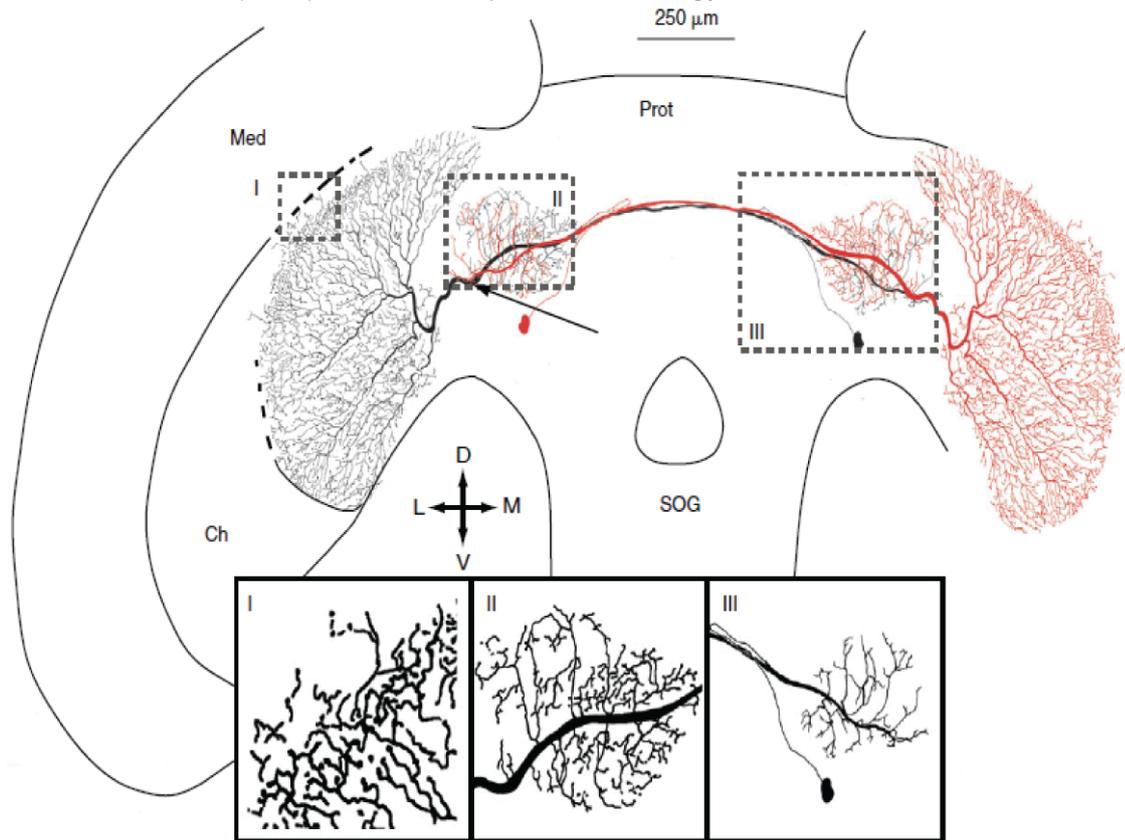


Figure 6-12. Morphology of CSTMD. Reproduced from Geurten et al (2007). Arrow indicates recording area from centrifugal STMD (black). Geurten et al suggested area III as the dendritic input (near cell soma and spike initiation zone). Area I and II were suggested to be likely output arborizations (beaded).

The primary input area for CSTMD (black) is in area III, on the contralateral side. The author's suggested that a mirror symmetric CSTMD (red) would be excited by the target on the ipsilateral side and then inhibit the CSTMD (black) under electrophysiological observation (interaction in input area III). This can explain the inhibition of spikes (in the initiation zone, area III) but does not explain my observation for a depolarisation (at the recording site, arrow), in response to the presence of a target on the ipsilateral side.

The question arises as to what is the benefit of this depolarisation? Perhaps the neuron is being primed (whilst the spiking is inhibited) for a target to translate into the contralateral hemisphere. Note though, that the mirror symmetric CSTMD would be weakly directional selective to the opposite direction of motion (progressive). Further electrophysiological investigation and modelling of these interactions is required.

Surprisingly, there has been limited investigation of STMDs and contrast polarity. In hoverfly, STMD responses to white targets have been observed, though they seemed to be only marginally beyond the spontaneous rate (Nordström et al., 2006).

The dark target selectivity that I observe in CSTMD is interesting from the modelling perspective. This suggests an underlying mechanism that includes a partial rectification to separate the dark contrast (an ‘off’ channel). With EMD-derived models, targets should provide responses largely independent of the sign of the contrast. To provide this polarity selectivity with an EMD model, a half-wave rectification is required to separate the contrast decrements prior to their multiplication. The ESTMD model is inherently contrast polarity selective, as it already includes this rectification stage.

CSTMD neural responses to the targets, in both images, are robust across many repeated trials, whilst the response to the other features appears more intermittent. The intermittent responsiveness to the bar-like features (columns in car park, blue arrows) is interesting, as the presentation of bars within a target height tuning scenario (on a white background) provides a continuous tuning curve, rather than intermittent spiking responses.

The ESTMD model was tuned for fly visual parameters such as spatial sampling, optical blur, size-selective inhibition and the value of the correlation delay time-constant. Therefore, an exact feature ‘matching’ (reliant on a spatiotemporal signature) between model and the dragonfly neuron was not expected. However, I had calculated that the differences in important attributes between the fly and dragonfly would cancel out in these experiments.

Generally, the CSTMD responses matched my expectations with robust target response, similar generation of false positives and the induced changes due to the image manipulation. Further aspects of the response can be readily attributed to the higher-order interactions previously described, such as a neural priming, temporal adaptation and inter-ocular inhibition.

These interactions are complex, and it will take time and further research to elucidate the mechanisms involved. For now though, it is clear that subtle manipulation of the images (determined by the ESTMD model) has accurately predicted dramatic changes in the observed electrophysiological responses. .

CHAPTER 7:

CONCLUSIONS

In this thesis I have presented a neuromorphic model (Chapters 2 and 3) for target discrimination in visual clutter based not on segregation of target motion from that of its background, but rather on its specificity for features. It is interesting to note that the stimuli for which biological STMD neurons are most sensitive appear to be relatively rare features in natural scenes allowing us to hone the thresholds for target discrimination (Chapters 4 and 5) so as to optimally reject the noise introduced by background clutter. Using optimal thresholds determined from modelling, my direct comparison between model predictions and STMD responses (Chapter 6) show that this model captures the performance of the biological system to a large degree.

Putative input elements and electrophysiological recording

Although my data make a compelling case for a relatively simple solution to a complex problem, much further work needs to be carried out before we conclude that the model I present is indeed the solution that has evolved in biology. Perhaps the biggest challenge facing future work to establish whether this is the case, lies in the tiny size of the retinotopically organized input elements in the insect visual pathway, particularly in the medulla. Although the results occupy only a few pages of text (Chapter 2), a major component of my research involved performing intracellular recordings from ‘on-off’ units within the medulla of the blowfly (*Calliphora stygia*), now referred to as the rectifying, transient cells (RTC). The ‘on-off’ units had previously been found to be temporally sluggish, an unlikely characteristic for an element within the target detection pathway. I examined the temporal responsiveness of the RTC to a more optimal stimulation (the doublet), examined the transient response, and discovered that these units had an equivalent frequency response to earlier neurons in the visual pathway. Generally, this approach may be useful in studying other highly adaptive, transient systems, where the steady-state responses are not functionally relevant. The discovery of these temporal characteristics, combined with the previously reported spatial properties, led to a newly hypothesised role in target discrimination.

While recording from other putative input elements proposed in my model would seem a natural next step in validating this model, neurons within the medulla of the fly are notoriously difficult to record from due to their small size and this then constrains the duration of the experiments. To obtain this RTC data set (independent adaptation and temporal responsiveness) I had to overcome many problems with electrophysiological techniques. This included finding the appropriate age window for the flies (4 to 7 days), finding a dissection method to access the medulla and development of a fly Ringer solution. In the end, recordings were of typically short duration, and the data presented are the end results of hundreds of hours or recording effort. These technical issues should be considered carefully by future researchers wishing to adopt a similar approach.

On-Off units

I developed a model of the ‘small target motion detector’ (STMD) neuron that included an RTC-type processing component and showed that it effectively discriminated targets from a cluttered surround. This finding revealed that within natural scenes, the target signature that biological STMDs are tuned to detect is quite rare. It is this rarity that allows the nonlinear filtering to aid in target discrimination. This is not to assert that RTCs or ‘on-off’ units *are* the elements along the pathway to the STMD neurons, but rather, that this research identified the suitability of this form of processing for the target detection task. Regardless of how much further computation is required in biological target discrimination, this mechanism is a good spatiotemporal pre-filter for the task, designed from known elements within the insect visual system.

As noted above, electrophysiological experiments from RTCs are a technically challenging task, though one worthy of further investigation. Now that we have examined some of the temporal properties, the next focus is to study the spatial characterisation of the RTC. This would involve experiments searching for any spatial antagonism and if found, the extent of these interactions. Due to the limited experimental duration when recording from RTCs, this requires an efficient visual stimulus apparatus designed for the task. Note that James and Osorio (1996) have successfully applied very rapid 2-dimensional white noise techniques to characterize spatial interactions within medulla columnar neurons in the locust, so application of this approach to the fly neurons may be informative.

Flicker versus motion and future modelling

The ESTMD model is not a true motion detector, but rather a temporal filter for a specific local flicker signature (with surround inhibitory interactions). This complex spatiotemporal signature is encountered with the motion of a dark target, but also with a localised black target appearance and then disappearance (flicker). An experiment with STMD neurons, with the visual stimulation of localised flicker targets, is unlikely to produce a neural response. Even STMD responses to two point ‘apparent motion’ are limited (Barnett, personal communications). It would seem that some form of higher-order signal integration, spatial or temporal, is required for robust STMD responses. This very robustness of the response to the features for which they are tuned makes investigation of STMDs with ‘elementary’ stimuli challenging.

STMDs can be direction selective, which from the perspective of this modelling, could emerge from either asymmetric inhibition or spatial facilitation during pooling of ESTMD units. Both of these aspects would require a directed target stimulus to drive the STMD neuron. Either of these directional aspects could be included in future modelling of the neuron.

Higher order receptive field structure

One of the most intriguing (indeed tantalizing) outcomes of my experiments to verify the model against biological STMD recordings (Chapter 6) was the observation that very subtle image manipulation of false positive features in natural scenes caused surprisingly large changes in the response of the neuron, not only at the location where those features had been obliterated, but at other regions within the larger receptive field. This highlights the complexity of the receptive field of the high-order neuron (the dragonfly CSTMD1) and a fundamental problem with my approach. My rationale for studying this particular neuron was that its large size created the possibility of the prolonged experiments required for quantitative analysis of responses to targets embedded in complex scenes. However, the trade-off is that (as revealed by my characterization of the receptive fields in chapter 6) the complex interaction of the inhibitory inputs from the contralateral visual hemisphere are clearly playing an important role in shaping the responses within these natural scenes. The CSTMD experiments revealed that the features predicted by our ESTMD model to be highly significant, when altered, had a dramatic effect on the output of the dragonfly CSTMD. These results revealed a complicated physiological system, with higher-order

complexities ensuring a complete interpretation is a challenging task. Nevertheless, these elaborate interactions did not mask the fact that we were able to invoke dramatic changes in neural response as predicted via the subtle image manipulations. This result is highly significant and should be investigated further.

An emergent hypothesis to explain my results is that release from this inhibition as targets cross the mid-line of the visual fields might ‘prime’ the global response to amplify the local excitation as the target enters the receptive field ‘hot spot’. This hypothesis can be addressed with further careful experiments using multiple target stimuli, but also highlights the need for additional modelling of the higher order receptive field properties. Hence, further modelling is required to include spatial integration over the receptive field. That is, a transition of my model from an ESTMD (i.e. a local stage of processing) to that for a true STMD. When integrating over the ESTMD outputs, it will be important to set the level of the threshold on each input appropriately. Too high, and target features within the receptive field are missed, and too low will cause spatial integration of ‘feature noise’. This is an area ripe for future investigation. The higher order aspects of the CSTMD can also be modelled, including the inter-ocular inhibitory interactions. These models will then form the basis for integration into pursuit algorithms, taking the target detection task into the realm of target tracking.

The model elaborations (5.2) have provided some theoretical considerations that can subserve the design of further electrophysiological investigation of higher order properties. The ESTMD model predicts robust target detection, without relative motion cues, as observed in the physiological system. However, a modulation of the STMD responses is expected if the system includes the mechanisms as described in the model elaborations. Careful design of the visual stimuli may be able to determine if this is the case.

In engineered systems, computer vision processing may be performed on moving platforms (e.g. UAVs). Approaches to feature detection mechanisms that are normalised to account for ego-motion are therefore a relevant area of investigation. Although preliminary, the algorithms as highlighted in this thesis may subserve this task.

The natural image inputs to the model can also be extended to include video input more akin to a $2D+t$ representation of visual ecology, including both rotatory and translatory components. The ability of the ESTMD (particularly with model elaborations) to discriminate targets embedded in these natural videos could be determined.

The scientific arena for the investigation of biological target detection is vast and further understanding will come from future endeavours in both electrophysiological and modelling efforts.

REFERENCES

- ADELSON, E. H. & BERGEN, J. R. (1985) Spatiotemporal energy models for the perception of motion. *Journal of the Optical Society of America A*, 2(2), 284-99.
- ARIEL, M. & ADOLPH, A. R. (1985) Neurotransmitter inputs to directionally sensitive turtle retinal ganglion cells. *Journal of Neurophysiology*, 54(5), 1123-43.
- ARNETT, D. W. (1971) Receptive field organization of units in the first optic ganglion of Diptera. *Science*, 173(4000), 929-931.
- ARNETT, D. W. (1972) Spatial and temporal integration properties of units in first optic ganglion of dipterans. *Journal of Neurophysiology*, 35(4), 429-444.
- ATICK, J. (1992) Could information theory provide an ecological theory of sensory processing? *Network: Computation in Neural Systems*, 3(2), 213-251.
- BARLOW, H. B. (1961) Possible principles underlying the transformation of sensory messages. IN ROSENBLITH, W. A. (Ed.) *Sensory Communication*. Cambridge, MIT Press.
- BARLOW, H. B. & LEVICK, W. R. (1965) The mechanism of directionally selective units in rabbit's retina. *Journal of Physiology*, 178(3), 477-504.
- BARNETT, P. D., NORDSTRÖM, K. & O'CARROLL, D. C. (2007) Retinotopic organization of small-field-target-detecting neurons in the insect visual system. *Current Biology*, 17(7), 569-578.
- BARRON, J. L., FLEET, D. J. & BEAUCHEMIN, S. S. (1994) Performance of optical flow techniques. *International Journal of Computer Vision*, 12(1), 43-77.
- BAUSENWEIN, B. & FISCHBACH, K. F. (1992) Activity labeling patterns in the medulla of *Drosophila-Melanogaster* caused by motion stimuli. *Cell and Tissue Research*, 270(1), 25-35.
- BEAUCHEMIN, S. S. & BARRON, J. L. (1995) The computation of optical flow. *ACM Computing Surveys (CSUR)*, 27(3), 433-466.
- BEERSMA, D. G. M., STAVENGA, D. G. & KUIPER, J. W. (1975) Organization of visual axes in the compound eye of the fly *Musca domestica L.* and behavioural consequences. *Journal of Comparative Physiology A - Neuroethology, Sensory, Neural, and Behavioral Physiology*, 102(4), 305-320.
- BEX, P. J. & MAKOUS, W. (2002) Spatial frequency, phase, and the contrast of natural images. *Journal of the Optical Society of America A*, 19(6), 1096-1106.
- BOEDDEKER, N., KERN, R. & EGELHAAF, M. (2003) Chasing a dummy target: smooth pursuit and velocity control in male blowflies. *Proceedings of the Royal Society of London Series B-Biological Sciences*, 270(1513), 393-399.
- BORST, A. & EGELHAAF, M. (1989) Principles of visual motion detection. *Trends in Neurosciences*, 12(8), 297-306.
- BORST, A., EGELHAAF, M. & HAAG, J. (1995) Mechanisms of dendritic integration underlying gain-control in fly motion-sensitive interneurons. *Journal of Computational Neuroscience*, 2(1), 5-18.
- BORST, A. & HAAG, J. (2002) Neural networks in the cockpit of the fly. *Journal of Comparative Physiology A - Neuroethology Sensory Neural and Behavioral Physiology*, 188(6), 419-437.

- BOUZERDOUM, A. (1993) The elementary movement detection mechanism in insect vision. *Philosophical Transactions of the Royal Society of London Series B - Biological Sciences*, 339(1290), 375-384.
- BOUZERDOUM, A. & PINTER, R. B. (1990) A shunting inhibitory motion detector that can account for the functional characteristics of fly motion-sensitive interneurons. *International Joint Conference on Neural Networks (IJCNN)*, 149-153.
- BRINKWORTH, R. S., MAH, E. L., GRAY, J. P. & D.C., O. C. (2008) Photoreceptor processing improves salience facilitating small target detection in cluttered scenes. *Journal of Vision*, 8(11), 8,1-17.
- BRINKWORTH, R. S. A., MAH, E. L. & O'CARROLL, D. C. (2006) Bioinspired pixel-wise adaptive imaging. *Proceedings of SPIE: Smart Structures, Devices, and Systems*, 3(6414), 641416,1-8.
- BRINKWORTH, R. S. A. & O'CARROLL, D. C. (2007) Biomimetic motion detection. *IEEE Proceedings, The International Conference on Intelligent Sensors, Sensor Networks and Information Processing*.
- BUCHNER, E. (1976) Elementary movement detectors in an insect visual system. *Biological Cybernetics*, 24(2), 85-101.
- BUCHNER, E. (1984) Behavioural analysis of spatial vision in insects. *Photoreception and Vision in Invertebrates*, 561-621.
- CAMPOS-ORTEGA, J. A. & STRAUSFELD, N. J. (1973) Synaptic connections of intrinsic cells and basket arborizations in the external plexiform layer of the fly's eye. *Brain Research*, 59119-36.
- CHOI, Y., ZAIJUN, P., KIM, S., KIM, T. & PARK, C. (2006) Salient motion information detection technique using weighted subtraction image and motion vector. *Proceedings of the 2006 International Conference on Hybrid Information Technology*, 1263-269.
- CLIFFORD, C. W., IBBOTSON, M. R. & LANGLEY, K. (1997) An adaptive Reichardt detector model of motion adaptation in insects and mammals. *Visual Neuroscience*, 14(4), 741-9.
- CLIFFORD, C. W. G. & IBBOTSON, M. R. (2002) Fundamental mechanisms of visual motion detection: models, cells and functions. *Progress in Neurobiology*, 68(6), 409-437.
- COLLETT, T. S. (1971) Visual neurones for tracking moving targets. *Nature*, 232(5306), 127-130.
- COLLETT, T. S. & KING, A. J. (1975) Vision during flight. IN HORRIDGE, G. A. (Ed.) *The compound eye and vision of insects*. Oxford, Clarendon Press.
- COLLETT, T. S. & LAND, M. F. (1975) Visual control of flight behaviour in the hoverfly, *Syritta pipiens*. *Journal of Comparative Physiology*, 991-66.
- COLLETT, T. S. & LAND, M. F. (1978) How hoverflies compute interception courses. *Journal of Comparative Physiology*, 125(3), 191-204.
- COOMBE, P. E., SRINIVASAN, M. V. & GUY, R. G. (1989) Are the Large Monopolar Cells of the insect lamina on the optomotor pathway. *Journal of Comparative Physiology A - Sensory Neural and Behavioral Physiology*, 166(1), 23-35.
- DEBEVEC, P. E. & MALIK, J. (1997) Recovering high dynamic range radiance maps from photographs. *SIGGRAPH*, 369-378.
- DELCOMYN, F. (1998) *Foundations of neurobiology*, New York, WH Freeman.

- DEVOE, R. D. & OCKLEFORD, E. M. (1976) Intracellular responses from cells of the medulla of the fly, *Calliphora erythrocephala*. *Biological Cybernetics*, 23(1), 13-24.
- DIETRICH, W. (1909) Die facettenaugen der Dipteren. *Zeitschrift für wissenschaftliche Zoologie*, (92), 465-539.
- DOUGLASS, J. K. & STRAUSFELD, N. J. (1995) Visual motion detection circuits in flies: peripheral motion computation by identified small-field retinotopic neurons. *Journal of Neuroscience*, 15(8), 5596-611.
- DOUGLASS, J. K. & STRAUSFELD, N. J. (1996) Visual motion-detection circuits in flies: parallel direction- and non-direction-sensitive pathways between the medulla and lobula plate. *Journal of Neuroscience*, 16(15), 4551-62.
- DOUGLASS, J. K. & STRAUSFELD, N. J. (2003) Anatomical organization of retinotopic motion-sensitive pathways in the optic lobes of flies. *Microscopy Research and Technique*, 62(2), 132-150.
- DROR, R. O., O'CARROLL, D. C. & LAUGHLIN, S. B. (2000) The role of natural image statistics in biological motion estimation. *Biologically Motivated Computer Vision, Lecture Notes in Computer Science*.
- EGELHAAF, M. (1985a) On the neuronal basis of figure-ground discrimination by relative motion in the visual-system of the fly .1. Behavioral constraints imposed on the neuronal network and the role of the optomotor system. *Biological Cybernetics*, 52(2), 123-140.
- EGELHAAF, M. (1985b) On the neuronal basis of figure-ground discrimination by relative motion in the visual-system of the fly .2. Figure-Detection cells, a new class of visual interneurons. *Biological Cybernetics*, 52(3), 195-209.
- EGELHAAF, M. (1985c) On the neuronal basis of figure-ground discrimination by relative motion in the visual-system of the fly. 3. Possible input circuitries and behavioral significance of the FD-cells. *Biological Cybernetics*, 52(4), 267-280.
- EGELHAAF, M. & BORST, A. (1993) A look into the cockpit of the fly - visual orientation, algorithms, and identified neurons. *Journal of Neuroscience*, 13(11), 4563-4574.
- EGELHAAF, M. & KERN, R. (2002) Vision in flying insects. *Current Opinion in Neurobiology*, 12(6), 699-706.
- FELSEN, G. & DAN, Y. (2005) A natural approach to studying vision. *Nature Neuroscience*, 8(12), 1643-1646.
- FERMI, G. & REICHARDT, W. (1963) Optomotor reactions of the fly, *Musca Domestica*. Dependence of the reaction on wavelength, velocity, contrast and median brightness of periodically moved stimulus patterns. *Kybernetik*, 18715-28.
- FIELD, D. J. (1987) Relations between the statistics of natural images and the response properties of cortical-cells. *Journal of the Optical Society of America A - Optics Image Science and Vision*, 4(12), 2379-2394.
- FRANCESCHINI, N., RIEHLE, A. & LE NESTOUR, A. (1989) Directionally selective motion detection by insect neurons. *Facets of Vision. Springer, Berlin, Heidelberg*, 360-390.
- GEIGER, G. & POGGIO, T. (1975) The orientation of flies towards visual patterns: On the search for the underlying functional interactions. *Biological Cybernetics*, 19(1), 39-54.
- GEURTEN, B. R. H., NORDSTROM, K., SPRAYBERRY, J. D. H., BOLZON, D. M. & O'CARROLL, D. C. (2007) Neural mechanisms underlying target detection

- in a dragonfly centrifugal neuron. *Journal of Experimental Biology*, 219(18), 3277-3284.
- GILBERT, C., PENISTEN, D. K. & DEVOE, R. D. (1991) Discrimination of visual motion from flicker by identified neurons in the medulla of the fleshfly *Sarcophaga bullata*. *Journal of Comparative Physiology A - Sensory Neural and Behavioral Physiology*, 168(6), 653-73.
- GILBERT, C. & STRAUSFELD, N. J. (1991) The functional organization of male-specific visual neurons in flies. *Journal of Comparative Physiology A - Sensory Neural and Behavioral Physiology*, 169(4), 395-411.
- HAAG, J., DENK, W. & BORST, A. (2004) Fly motion vision is based on Reichardt detectors regardless of the signal-to-noise ratio. *Proceedings of the National Academy of Sciences of the United States of America*, 101(46), 16333-16338.
- HANLEY, J. A. & MCNEIL, B. J. (1982) The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology*, 143(1), 29-36.
- HARDIE, R. C. (1985) Functional organization of the fly retina. *Progress in Sensory Physiology*, 51-79.
- HARDIE, R. C. (1987) Is histamine a neurotransmitter in insect photoreceptors? *Journal of Comparative Physiology A - Sensory Neural and Behavioral Physiology*, 161(2), 201-13.
- HARDIE, R. C. (2001) Phototransduction in *Drosophila melanogaster*. *Journal of Experimental Biology*, 204(20), 3403-3409.
- HARDIE, R. C., FRANCESCHINI, N., RIBI, W. & KIRSCHFELD, K. (1981) Distribution and properties of sex-specific photoreceptors in the fly *Musca domestica*. *Journal of Comparative Physiology A - Sensory Neural and Behavioral Physiology*, 145(2), 139-152.
- HARDIE, R. C. & WECKSTRÖM, M. (1990) Three classes of potassium channels in large monopolar cells of the blowfly *Calliphora vicina*. *Journal of Comparative Physiology A - Sensory Neural and Behavioral Physiology*, 167723-736.
- HARRIS, C. & STEPHENS, M. (1988) A combined corner and edge detector. *Alvey Vision Conference*, 1550.
- HARRIS, R. A., O'CARROLL, D. C. & LAUGHLIN, S. B. (1999) Adaptation and the temporal delay filter of fly motion detectors. *Vision Research*, 39(16), 2603-2613.
- HASSENSTEIN, B. & REICHARDT, W. (1956) Systemtheoretische analyse der zeitreihenfolgen und vorzeichenauswertung bei der bewegungsperzeption des rüsselkäfers *Chlorophanus*. *Z. Naturforsch.*, (11b), 513-524.
- HATEREN, J. H., HARDIE, R. C., RUDOLPH, A., LAUGHLIN, S. B. & STAVENGA, D. G. (1989) The bright zone, a specialized dorsal eye region in the male blowfly *Chrysomya megacephala*. *Journal of Comparative Physiology A: Sensory, Neural, and Behavioral Physiology*, 164(3), 297-308.
- HATSOPOULOS, N., GABBIANI, F. & LAURENT, G. (1995) Elementary computation of object approach by a wide-field visual neuron. *Science*, 270(5238), 1000.
- HAUSEN, K. (1982) Motion sensitive interneurons in the optomotor system of the fly .1. The Horizontal Cells - Structure and signals. *Biological Cybernetics*, 45(2), 143-156.
- HAUSEN, K. (1993) Decoding of retinal image flow in insects. *Reviews of Oculomotor Research*, 5203-35.

- HAUSEN, K. & STRAUSFELD, N. J. (1980) Sexually dimorphic interneuron arrangements in the fly visual system. *Proceedings of the Royal Society of London Series B-Biological Sciences*, 208(1170).
- HIGGINS, C. M., DOUGLASS, J. K. & STRAUSFELD, N. J. (2004) The computational basis of an identified neuronal circuit for elementary motion detection in dipterous insects. *Visual Neuroscience*, 21(4), 567-586.
- HIGGINS, C. M. & PANT, V. (2004) An elaborated model of fly small-target tracking. *Biological Cybernetics*, 91(6), 417-428.
- HOLT, G. R. & KOCH, C. (1997) Shunting inhibition does not have a divisive effect on firing rates. *Neural Computation*, 9(5), 1001-1013.
- HORNSTEIN, E. P., O'CARROLL, D. C., ANDERSON, J. C. & LAUGHLIN, S. B. (2000) Sexual dimorphism matches photoreceptor performance to behavioural requirements. *Proceedings of the Royal Society of London Series B-Biological Sciences*, 267(1457), 2111-2117.
- HORRIDGE, G. A. (1978) The separation of visual axes in apposition compound eyes. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences (1934-1990)*, 285(1003), 1-59.
- IBBOTSON, M. R., MARK, R. F. & MADDESS, T. L. (1994) Spatiotemporal response properties of direction-selective neurons in the nucleus of the optic tract and dorsal terminal nucleus of the wallaby, *Macropus eugenii*. *Journal of Neurophysiology*, 72(6), 2927-2943.
- JAMES, A. C. (1990) White-noise studies in the fly lamina [PhD Thesis]. Canberra, Australian National University.
- JAMES, A. C. (1992) Nonlinear operator network models of processing in the fly lamina. IN B, N. (Ed.) *Nonlinear Vision*. Boca Raton, FL, CRC.
- JAMES, A. C. & OSORIO, D. (1996) Characterisation of columnar neurons and visual signal processing in the medulla of the locust optic lobe by system identification techniques. *Journal of Comparative Physiology A: Sensory, Neural, and Behavioral Physiology*, 178(2), 183-199.
- JANSONIUS, N. M. & VAN HATEREN, J. H. (1991) Fast temporal adaptation of On-Off units in the 1st optic chiasm of the blowfly. *Journal of Comparative Physiology A - Sensory Neural and Behavioral Physiology*, 168(6), 631-637.
- JANSONIUS, N. M. & VAN HATEREN, J. H. (1993a) On-Off units in the 1st optic chiasm of the blowfly. 2. Spatial properties. *Journal of Comparative Physiology A - Sensory Neural and Behavioral Physiology*, 172(4), 467-471.
- JANSONIUS, N. M. & VAN HATEREN, J. H. (1993b) On spiking units in the 1st optic chiasm of the blowfly .3. The sustaining unit. *Journal of Comparative Physiology A - Sensory Neural and Behavioral Physiology*, 173(2), 187-192.
- JARVILEHTO, M. & ZETTLER, F. (1971) Localized intracellular potentials from pre- and post-synaptic components in the external plexiform layer of an insect retina. *Zeitschrift fur Vergleichende Physiologie*, 75422-440.
- JÄRVILEHTO, M. & ZETTLER, F. (1970) Micro-localisation of lamina-located visual cell activities in the compound eye of the blowfly *Calliphora*. *Journal of Comparative Physiology A - Sensory, Neural, and Behavioral Physiology*, 69(1), 134-138.
- JUUSOLA, M., KOUVALAINEN, E., JARVILEHTO, M. & WECKSTROM, M. (1994) Contrast gain, signal-to-noise ratio, and linearity in light-adapted blowfly photoreceptors. *Journal of General Physiology*, 104(3), 593-621.

- JUUSOLA, M., UUSITALO, R. O. & WECKSTROM, M. (1995) Transfer of graded potentials at the photoreceptor interneuron synapse. *Journal of General Physiology*, 105(1), 117-148.
- KELLER, A. (2002) Genetic intervention in sensory systems of a fly [PhD Thesis]. Wurzburg, Bayerischen Julius-Maximilians-Universität.
- KIMMERLE, B. & EGELHAAF, M. (2000) Performance of fly visual interneurons during object fixation. *Journal of Neuroscience*, 20(16), 6256-6266.
- KIRSCHFELD, K. & WENK, P. (1976) The dorsal compound eye of simuliid flies: an eye specialized for the detection of small, rapidly moving objects. *Z Naturforsch [C]*, 31(11-12), 764-5.
- KOCH, C. (2004) *Biophysics of Computation: Information Processing in Single Neurons (Computational Neuroscience Series) [21.1.1]*, Oxford University Press, Inc. New York, NY, USA.
- KRAPP, H. G. & HENGSTENBERG, R. (1996) Estimation of self-motion by optic flow processing in single visual interneurons. *Nature*, 384(6608), 463-466.
- LAND, M. F. (1981) Optics and vision in invertebrates. *Handbook of Sensory Physiology*, 7471-592.
- LAND, M. F. (1993) Chasing and pursuit in the Dolichopodid fly *Poecilobothrus Nobilitatus*. *Journal of Comparative Physiology A - Sensory Neural and Behavioral Physiology*, 173(5), 605-613.
- LAND, M. F. (1997) Visual acuity in insects. *Annual Review of Entomology*, 42147-177.
- LAND, M. F. & COLLETT, T. S. (1974) Chasing behaviour of house flies (*Fannia canicularis*). *Journal of Comparative Physiology*, 89331-357.
- LAND, M. F. & ECKERT, H. (1985) Maps of the acute zones of fly eyes. *Journal of Comparative Physiology A - Sensory Neural and Behavioral Physiology*, 156(4), 525-538.
- LAUGHLIN, S. (1981a) A simple coding procedure enhances a neuron's information capacity. *Z Naturforsch [C]*, 36(9-10), 910-2.
- LAUGHLIN, S. (1984) The roles of parallel channels in early visual processing by the arthropod compound eye. *Photoreception and vision in invertebrates*. Plenum, New York, 457-481.
- LAUGHLIN, S. B. (1981b) Neural principles in the peripheral visual systems of invertebrates. IN AUTRUM, H. (Ed.) *Handbook of Sensory Physiology*. Berlin, Springer.
- LAUGHLIN, S. B. (1994) Matching coding, circuits, cells, and molecules to signals - general principles of retinal design in the fly's eye. *Progress in Retinal and Eye Research*, 13(1), 165-196.
- LAUGHLIN, S. B., HOWARD, J. & BLAKESLEE, B. (1987) Synaptic limitations to contrast coding in the retina of the blowfly *Calliphora*. *Proceedings of the Royal Society B: Biological Sciences*, 231(1265), 437-67.
- LAUGHLIN, S. B. & WECKSTROM, M. (1993) Fast and slow photoreceptors - a comparative- study of the functional diversity of coding and conductances in the Diptera. *Journal of Comparative Physiology A - Sensory Neural and Behavioral Physiology*, 172(5), 593-609.
- LIMB, J. O. & MURPHY, J. A. (1975) Estimating the velocity of moving images in television signals. *Comput. Graphics Image Processing*, 4311-327.
- MAH, E. L., BRINKWORTH, R. S. & O'CARROLL, D. C. (2008a) An elaborated electronic prototype of a biological photoreceptor. *Biological Cybernetics*, 98357-369.

- MAH, E. L., BRINKWORTH, R. S. & O'CARROLL, D. C. (2008b) Implementation of an elaborated neuromorphic model of a biological photoreceptor. *Biological Cybernetics*.
- MAH, E. L., BRINKWORTH, R. S. A. & O'CARROLL, D. (2005) Bio-inspired analog circuitry model of insect photoreceptor cells. *Proceedings of SPIE*, 6036603613.
- MCNEIL, B. J. & HANLEY, J. A. (1984) Statistical approaches to the analysis of Receiver Operating Characteristic (ROC) curves. *Medical Decision Making*, 4(2), 137.
- MICHELSON, A. (1927) *Studies in Optics*. Chicago: Chicago University Press (Reissued 1962 Phoenix Science Series).
- MIMURA, K. (1972) Neural mechanisms subserving directional selectivity of movement in the optic lobe of the fly. *Journal of Comparative Physiology*, 80409-437.
- MIZUTANI, A., CHAHL, J. S. & SRINIVASAN, M. V. (2003) Insect behaviour: Motion camouflage in dragonflies. *Nature*, 423(6940), 604-604.
- MORAVEC, H. P. (1977) *Towards automatic visual obstacle avoidance*. USA.
- NILSSON, D. E. (1989) Optics and evolution of the compound eye. *Facets of vision*. Springer, Berlin Heidelberg New York, 30-73.
- NORDSTRÖM, K., BARNETT, P. D., DE MIGUEL, I. M. M., BRINKWORTH, R. S. A. & O'CARROLL, D. C. (2008) Sexual dimorphism in the hoverfly motion vision pathway. *Current Biology*, 18(9), 661-667.
- NORDSTRÖM, K., BARNETT, P. D. & O'CARROLL, D. C. (2006) Insect detection of small targets moving in visual clutter. *PLoS Biology*, 4(3), 378-386.
- NORDSTRÖM, K. & O'CARROLL, D. C. (2006) Small object detection neurons in female hoverflies. *Proceedings of the Royal Society B: Biological Sciences*, 273(1591), 1211-1216.
- O'CARROLL, D. (1993) Feature-detecting neurons in dragonflies. *Nature*, 362(6420), 541-543.
- O'CARROLL, D. C., LAUGHLIN, S. B., BIDWELL, N. J. & HARRIS, R. A. (1997) Spatio-temporal properties of motion detectors matched to low image velocities in hovering insects. *Vision Research*, 37(23), 3427-3439.
- O'CARROLL, D. C., OSORIO, D., JAMES, A. C. & BUSH, T. (1992) Local feedback mediated via amacrine cells in the insect optic lobe. *Journal of Comparative Physiology a-Sensory Neural and Behavioral Physiology*, 171(4), 447-455.
- O'SHEA, M. & WILLIAMS, J. L. D. (1974) The anatomy and output connection of a locust visual interneurone; the lobular giant movement detector (LGMD) neurone. *Journal of Comparative Physiology A - Sensory, Neural, and Behavioral Physiology*, 91(3), 257-266.
- OGMEN, H. & GAGNE, S. (1988) Short-range motion detection in the insect visual system. *Neural Networks*, 1519.
- OGMEN, H. & GAGNE, S. (1990) Neural models for sustained and On-Off units of insect lamina. *Biological Cybernetics*, 63(1), 51-60.
- OLBERG, R. M. (1981) Object-movement and self-movement detectors in the ventral nerve cord of the dragonfly. *Journal of Comparative Physiology*, 141(3), 327-334.
- OLBERG, R. M. (1986) Identified target-selective visual interneurons descending from the dragonfly brain. *Journal of Comparative Physiology A - Sensory Neural and Behavioral Physiology*, 159(6), 827-840.

- OLBERG, R. M., WORTHINGTON, A. H. & VENATOR, K. R. (2000) Prey pursuit and interception in dragonflies. *Journal of Comparative Physiology A - Sensory Neural and Behavioral Physiology*, 186(2), 155-162.
- OLSHAUSEN, B. A. & FIELD, D. J. (1996) Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *Nature*, 381(6583), 607-609.
- OSORIO, D. (1986) Directionally selective cells in the locust medulla. *Journal of Comparative Physiology A - Sensory Neural and Behavioral Physiology*, 159(6), 841-7.
- OSORIO, D. (1987) The temporal properties of nonlinear, transient cells in the locust medulla. *Journal of Comparative Physiology A - Sensory Neural and Behavioral Physiology*, 161(3), 431-440.
- OSORIO, D. (1991) Mechanisms of early visual processing in the medulla of the locust optic lobe - How self-inhibition, spatial-pooling, and signal rectification contribute to the properties of transient cells. *Visual Neuroscience*, 7(4), 345-355.
- PALKA, J. (1969) Discrimination between movements of eye and object by visual interneurons of crickets. *Journal of Experimental Biology*, 50(3), 723-732.
- PEARSON, K. G., HEITLER, W. J. & STEEVES, J. D. (1980) Triggering of locust jump by multimodal inhibitory interneurons. *Journal of Neurophysiology*, 43(2), 257-278.
- POTTERS, M. & BIALEK, W. (1994) Statistical mechanics and visual signal processing. *Journal of Physiology France*, (4), 1755-1775.
- REICHARDT, W. (1961) Autocorrelation, a principle for the evaluation of sensory information by the central nervous system. IN ROSENBLITH, W. A. (Ed.) *Sensory Communication*. New York, Wiley.
- REICHARDT, W. & POGGIO, T. (1979) Figure-ground discrimination by relative movement in the visual-system of the fly .1. Experimental results. *Biological Cybernetics*, 35(2), 81-100.
- REICHARDT, W., POGGIO, T. & HAUSEN, K. (1983) Figure-ground discrimination by relative movement in the visual-system of the fly .2. Towards the neural circuitry. *Biological Cybernetics*, 461-30.
- RIEKE, F., BODNAR, D. A. & BIALEK, W. (1995) Naturalistic stimuli increase the rate and efficiency of information transmission by primary auditory afferents. *Proceedings of the Royal Society B: Biological Sciences*, 262(1365), 259-265.
- RIND, F. C. & BRAMWELL, D. I. (1996) Neural network based on the input organization of an identified neuron signaling impending collision. *Journal of Neurophysiology*, 75(3), 967-985.
- RIND, F. C. & SIMMONS, P. J. (1992) Orthopteran DCMD neuron: a reevaluation of responses to moving objects. I. Selective responses to approaching objects. *Journal of Neurophysiology*, 68(5), 1654-1666.
- ROWELL, C. H., O'SHEA, M. & WILLIAMS, J. L. (1977) The neuronal basis of a sensory analyser, the acridid movement detector system. 4. The preference for small field stimuli. *Journal of Experimental Biology*, 68(1), 157-185.
- RUDERMAN, D. L. (1994) The statistics of natural images. *Network-Computation in Neural Systems*, 5(4), 517-548.
- RUST, N. C. & MOVSHON, J. A. (2005) In praise of artifice. *Nature Neuroscience*, 8(12), 1647-1650.

- SARIKAYA, M., WANG, W. X. & OGMEN, H. (1998) Neural network model of on-off units in the fly visual system: simulations of dynamic behavior. *Biological Cybernetics*, 78(5), 399-412.
- SHAW, S. R. (1984) Early visual processing in insects. *Journal of Experimental Biology*, 112(1), 225-251.
- SHI, J. & TOMASI, C. (1994) Good features to track. *Computer Vision and Pattern Recognition, 1994. Proceedings CVPR'94., 1994 IEEE Computer Society Conference on*, 593-600.
- SHOEMAKER, P. A., O'CARROLL, D. C. & STRAW, A. D. (2005) Velocity constancy and models for wide-field visual motion detection in insects. *Biological Cybernetics*, 93(4), 275-287.
- SIMONCELLI, E. P. & OLSHAUSEN, B. A. (2001) Natural image statistics and neural representation. *Annual Review of Neuroscience*, 24, 1193-1216.
- SINGLE, S. & BORST, A. (1998) Dendritic integration and its role in computing image velocity. *Science*, 281(5384), 1848-50.
- SINGLE, S., HAAG, J. & BORST, A. (1997) Dendritic computation of direction selectivity and gain control in visual interneurons. *Journal of Neuroscience*, 17(16), 6023-6030.
- SMITH, S. M. & BRADY, J. M. (1997) SUSAN—A new approach to low level image processing. *International Journal of Computer Vision*, 23(1), 45-78.
- SOBEY, P. & SRINIVASAN, M. V. (1991) Measurement of optical flow by a generalized gradient scheme. *Journal of the Optical Society of America A*, 8(9), 1488-1498.
- SRINIVASAN, M. V. (1990) Generalized gradient schemes for the measurement of two-dimensional image motion. *Biological Cybernetics*, 63(6), 421-31.
- SRINIVASAN, M. V. & BERNARD, G. D. (1977) The pursuit response of the housefly and its interaction with the optomotor response. *Journal of Comparative Physiology A - Sensory, Neural, and Behavioral Physiology*, 115(1), 101-117.
- SRINIVASAN, M. V. & GUY, R. G. (1990) Spectral properties of movement perception in the dronefly *Eristalis*. *Journal of Comparative Physiology A: Sensory, Neural, and Behavioral Physiology*, 166(3), 287-295.
- SRINIVASAN, M. V., LAUGHLIN, S. B. & DUBS, A. (1982) Predictive coding - a fresh view of inhibition in the retina. *Proceedings of the Royal Society of London Series B-Biological Sciences*, 216(1205), 427-459.
- SRINIVASAN, M. V., LEHRER, M. & HORRIDGE, G. A. (1990a) Visual figure-ground discrimination in the honeybee: The role of motion parallax at boundaries. *Proceedings of the Royal Society of London. Series B, Biological Sciences*, 238(1293), 331-350.
- SRINIVASAN, M. V., LEHRER, M., KIRCHNER, W. H. & ZHANG, S. W. (1991) Range perception through apparent image speed in freely flying honeybees. *Visual Neuroscience*, 6(5), 519-35.
- SRINIVASAN, M. V., PINTER, R. B. & OSORIO, D. (1990b) Matched filtering in the visual system of the fly - Large Monopolar Cells of the lamina are optimized to detect moving edges and blobs. *Proceedings of the Royal Society of London Series B-Biological Sciences*, 240(1298), 279-293.
- STAVENGA, D. G. (2003) Angular and spectral sensitivity of fly photoreceptors. I. Integrated facet lens and rhabdomere optics. *Journal of Comparative Physiology A - Neuroethology Sensory Neural and Behavioral Physiology*, 189(1), 1-17.

- STRAUSFELD, N. J. (1976) Some quantitative aspects of the fly's brain. *Atlas of an Insect Brain*. Heidelberg, Germany, Springer-Verlag.
- STRAUSFELD, N. J. (1989) Beneath the compound eye: neuroanatomical analysis and physiological correlates in the study of insect vision. *Facets of vision*. Springer, Berlin Heidelberg New York, 317–359.
- STRAUSFELD, N. J. (1991) Structural Organization of Male-Specific Visual Neurons in Calliphorid Optic Lobes. *Journal of Comparative Physiology a-Sensory Neural and Behavioral Physiology*, 169(4), 379-393.
- STRAUSFELD, N. J. & CAMPOS-ORTEGA, J. A. (1977) Vision in insects: pathways possibly underlying neural adaptation and lateral inhibition. *Science*, 195(4281), 894-7.
- STRAUSFELD, N. J. & NÄSSEL, D. R. (1980) Neuroarchitecture of brain regions that subserve the compound eyes of crustacea and insects. IN AUTRUM, H. (Ed.) *Handbook of sensory physiology*. New York, Springer-Verlag Berlin.
- STRAUSFELD, N. J. & NÄSSEL, D. R. (1981) Neuroarchitectures serving compound eyes of crustacea and insects. IN AUTRUM, H. (Ed.) *Handbook of Sensory Physiology*. New York, Springer-Verlag Berlin.
- STRAW, A. D., WARRANT, E. J. & O'CARROLL, D. C. (2006) A 'bright zone' in male hoverfly (*Eristalis tenax*) eyes and associated faster motion detection and increased contrast sensitivity. *Journal of Experimental Biology*, 209(21), 4339-4354.
- TANG, P., GAO, L. & LIU, Z. (2007) Salient moving object detection using stochastic approach filtering. *Image and Graphics, 2007. ICIG 2007. Fourth International Conference on*, 530-535.
- TOLHURST, D. J., TADMOR, Y. & CHAO, T. (1992) Amplitude spectra of natural images. *Ophthalmic and Physiological Optics*, 12(2), 229-232.
- TORRE, V. & POGGIO, T. (1978) A synaptic mechanism possibly underlying directional selectivity to motion. *Proceedings of the Royal Society of London. Series B, Biological Sciences (1934-1990)*, 202(1148), 409-416.
- VAN HATEREN, J. H. (1992a) Real and optimal neural images in early vision. *Nature*, 360(6399), 68-70.
- VAN HATEREN, J. H. (1992b) Theoretical predictions of spatiotemporal receptive-fields of fly LMCs, and experimental validation. *Journal of Comparative Physiology A - Sensory Neural and Behavioral Physiology*, 171(2), 157-170.
- VAN HATEREN, J. H. (1992c) A theory of maximizing sensory information. *Biological Cybernetics*, 68(1), 23-29.
- VAN HATEREN, J. H. (1993) 3 Modes of spatiotemporal preprocessing by eyes. *Journal of Comparative Physiology A - Sensory Neural and Behavioral Physiology*, 172(5), 583-591.
- VAN HATEREN, J. H. (1997) Processing of natural time series of intensities by the visual system of the blowfly. *Vision Research*, 37(23), 3407-3416.
- VAN HATEREN, J. H. (1998) Independent component filters of natural images compared with simple cells in primary visual cortex. *Proceedings of the Royal Society B: Biological Sciences*, 265(1394), 359-366.
- VAN HATEREN, J. H. & SNIPPE, H. P. (2001) Information theoretical evaluation of parametric models of gain control in blowfly photoreceptor cells. *Vision Research*, 41(14), 1851-1865.
- VAN SANTEN, J. P. H. & SPERLING, G. (1983) A new class of models of motion-sensitive units in human vision. *16th Annual Meeting of the Society for Mathematical Psychology, Boulder, Colorado*.

- VAN SANTEN, J. P. H. & SPERLING, G. (1985) Elaborated Reichardt detectors. *Journal of the Optical Society of America A - Optics Image Science and Vision*, 2(2), 300-321.
- WACHENFELD, A. (1994) Elektrophysiologische untersuchungen und funktionelle charakterisierung männchenspezifischer visueller interneurone de schmeißfliege *Calliphore erythrocephala* [PhD Thesis]. Köln, Universität zu Köln.
- WAGNER, H. (1986a) Flight performance and visual control of flight of the free-flying housefly (*Musca-Domestica L*) .1. Organization of the flight motor. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences*, 312(1158), 527-551.
- WAGNER, H. (1986b) Flight performance and visual control of flight of the free-flying housefly (*Musca domestica L*) 3. Interactions between angular movement induced by widefield and smallfield stimuli. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences*, 312(1158), 581-595.
- WARZECHA, A. K., EGELHAAF, M. & BORST, A. (1993) Neural circuit tuning fly visual interneurons to motion of small objects .1. Dissection of the circuit by pharmacological and photoinactivation techniques. *Journal of Neurophysiology*, 69(2), 329-339.
- WEHRHAHN, C. (1979) Sex-specific differences in the chasing behaviour of houseflies (musca). *Biological Cybernetics*, 32(239), 239-241.
- WEHRHAHN, C., POGGIO, T. & BULTHOFF, H. (1982) Tracking and chasing in houseflies (*Musca*) - an analysis of 3-D flight trajectories. *Biological Cybernetics*, 45(2), 123-130.
- WIEDERMAN, S., SHOEMAKER, P. A. & O'CARROLL, D. C. (2007) Biologically inspired small target detection mechanisms. *IEEE Proceedings, The International Conference on Intelligent Sensors, Sensor Networks and Information Processing*.
- WIEDERMAN, S. D., SHOEMAKER, P. A. & O'CARROLL, D. C. (2008) A model for the detection of moving targets in visual clutter inspired by insect physiology. *PLoS ONE*, 3(7), e2784.
- WIXSON, L. (2000) Detecting salient motion by accumulating directionally-consistent flow. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(8), 774-780.
- WOLF-OBERHOLLENZER, F. & KIRSCHFELD, K. (1994) Motion sensitivity in the nucleus of the basal optic root of the pigeon. *Journal of Neurophysiology*, 71(4), 1559-1573.
- YAMAGUCHI, S., WOLF, R., DESPLAN, C. & HEISENBERG, M. (2008) Motion vision is independent of color in *Drosophila*. *Proceedings of the National Academy of Sciences*, 105(12), 4910.
- YARFITZ, S. & HURLEY, J. B. (1994) Transduction mechanisms of vertebrate and invertebrate photoreceptors. *Journal of Biological Chemistry*, 269(20), 14329-32.
- ZANKER, J. M., SRINIVASAN, M. V. & EGELHAAF, M. (1999) Speed tuning in elementary motion detectors of the correlation type. *Biological Cybernetics*, 80(2), 109-16.
- ZEIL, J. (1983) Sexual dimorphism in the visual system of flies: The compound eyes and neural superposition in bionidae (Diptera). *Journal of Comparative Physiology A - Neuroethology, Sensory, Neural, and Behavioral Physiology*, 150(3), 379-393.

ZHENG, L., DE POLAVIEJA, G. G., WOLFRAM, V., ASYALI, M. H., HARDIE, R. C. & JUUSOLA, M. (2005) Feedback network controls photoreceptor output at the layer of first visual synapses in *Drosophila*. *Journal of General Physiology*, 127(5), 495-510.