A Neuromorphic Model For A Robust, Adaptive Photoreceptor Reduces Variability In Correlation Based Motion Detectors

David C. O’Carroll, Paul D. Barnett, Eng Leng Mah, Karin Nordström and Russell S. A. Brinkworth
Discipline of Physiology, School of Molecular & Biomedical Science and Computational Neuroscience Research Cluster
The University of Adelaide, SA 5005, Australia

ABSTRACT

We present here a parametric model for motion detection based on a correlational elementary motion detector (EMD) and inspired by our analysis of responses of neurons in the motion detection pathway of flying insects. This model incorporates a biomimetic photoreceptor stage that fully accounts for the non-linear adaptive normalization of contrast observed in fly photoreceptors. We find that inclusion of this front-end leads to a substantial improvement in performance compared with a basic EMD model. Our model lends itself to elaboration into analog electronic hardware, including neuromorphic analog VLSI. We have developed an initial implementation of the photoreceptor model and a single EMD using discrete electronic components. We have tested both the hardware and digital simulations of elaborated EMD arrays using high dynamic range (HDR) panoramic scenes derived from nature. Our data confirm that this photoreceptor model is robust enough to have a variety of applications and should be used as a front end wherever wide-field velocity information is of value (e.g. in optical flow analysis).

INTRODUCTION

The correlation family of motion detector models (including so called motion energy models) have been used to explain direction selective motion detection in a wide variety of insects, birds and mammals, including humans[1]. These models involve a non-linear correlation of adjacent spatial samples, with an asymmetric delay filter giving rise to direction selective responses within a local elementary motion detector (EMD)[2]. Arrays of EMDs are then summed to provide measurements of wide-field optical flow or motion of specific targets. This type of motion detector has been shown to have significant advantages over alternatives (e.g. gradient detection schemes) where detector noise is problematic (e.g. at low contrasts or luminance)[3]. Despite their ubiquity in nature, and several attempts to develop artificial motion detectors using neuromorphic principles (e.g. devices in analog VLSI)[4, 5], Such models have had limited success when applied to real-world artificial motion detection tasks. This is due in large part to the inherent sensitivity of correlation-based EMDs to contrast and spatial structure of local features within moving scenes. This leads to ambiguity in the local EMD response as a function of velocity, a phenomenon we term ‘pattern noise’[6, 7]. While this local ambiguity may be problematic, much evidence suggests that such motion detectors are used by insects to solve numerous problems of great interest to artificial vision.

Insects are able to analyze the velocity of optical flow, and to use this information in a variety of complex navigational tasks[8]. In recent years, we have studied the electrophysiology of wide field neurons in the fly brain which display surprisingly robust responses to velocity of optical flow in a wide variety of natural scenes that ought to produce confounded responses from ‘classical’ correlation EMDs[9]. We have subsequently studied and modeled additional non-linear processing stages in the insect motion pathway such as local, non-linear adaptation (local automatic gain control), that improve the robustness of artificial EMDs, but have not hitherto been able to match the precision of the biological system[10].

Our previous modeling efforts include front end (photodetector) components that capture some key elements of biological visual processing such as bandpass spatial and temporal filtering and logarithmic coding of luminance, but many of the filters are linear and thus simplified compared with those involved in biological photodetection. Recently, van Hateren and colleagues[11] developed a parametric model for early visual processing by insects that, when tested against the system it is designed to mimic, accounted thoroughly for much of the non-linear coding (including contrast normalization) of natural image sequences, particularly where those scenes have high contrast. We hypothesized that inclusion of such a model as a front end to an elaborated EMD model will provide a more robust mechanism for velocity analysis of scenes with variable spatial structure and contrast – the very conditions that generate ambiguous outputs using ‘traditional’ EMD models causing them to fail.

In this paper, we present a model for motion detection based on a correlational EMD (Fig. 1B) and inspired by our analysis of responses of neurons in the motion detection pathway of flying insects. This model incorporates a biomimetic photoreceptor stage based on the van Hateren and Snippe model for fly photoreceptors with some additional elaborations/tuning[12] is shown in Fig. 1A. We find that the inclusion of this front-end, in combination with additional elaborations inspired directly by our own experiments on.
adaptation of insect motion detectors, leads to a substantial improvement in performance compared with a basic EMD model. Our model lends itself to elaboration into analog electronic hardware, including neuromorphic analogue VLSI[13]. We have developed an initial implementation of the photoreceptor model and a single EMD using discrete electronic components. We have tested both the hardware and software simulations of EMD arrays using high dynamic range (HDR) panoramic scenes derived from nature. We have also tested the same scenes presented directly to directional motion-sensitive HS neurons in the motion pathway of the hoverfly Eristalis. The model provides a better fit to the biological response than our earlier attempts to model this system. Our data confirm that this photoreceptor model is robust enough to have a variety of applications and should be used wherever wide-field velocity information is of value (e.g. in optical flow analysis).

**EXPERIMENTATION**

**IMAGE CAPTURE**

We used a Nikon D-70 digital camera to obtain 13 HDR panoramic images from a variety of man-made and natural locations (Fig. 2). Locations were selected to represent a range of luminance and contrast conditions. Each panorama was obtained using a series of 12 overlapping panels saved in NEF (raw) format. In order to capture components of the scenes that exceeded the dynamic range of the camera sensor, each panel was captured at a range of exposure levels. For each exposure level over-saturated pixels were discarded and local luminance was then established using calibration of the luminance/value (gamma) curve for the camera[14] and then converted to floating point format using custom software written in LabView®.

Table I shows the contrast and other image statistics for the 13 images illustrated in Fig. 2. We considered both global image statistics and those with respect to the motion detector model (and the biological system it is intended to mimic). Since HS neurons are selective for horizontal motion, we evaluated both global and row contrast.
By both contrast measures there was a large contrast range among the 13 images. Globally, the images span an almost 6-fold range with the highest global contrast an indoor car-park scene (image D), while an outdoor scene relatively devoid of foreground landmarks or background relief had the lowest contrast (image B). Similar trends held when row contrast was evaluated. Since image B had the least obvious vertically aligned features, it was expected to produce the weakest response from correlation motion detectors selective for horizontal motion.

**BASIC EMD MODEL**

To highlight the problem with the basic EMD model for analyzing the velocity of scenes of variable contrast, we constructed a simple EMD model that used the raw luminance values of the images, divided by their mean value, as inputs. Because the images are circular, we were able to animate them continuously and record the output of an EMD in steady state for each image row. Because the fly that we were attempting to model has optics that blur the high spatial frequency content of the scene, we pre-blurred each image using a Gaussian blur of 1.4 degree half-width [15, 16].

The input pixels to the model were then sampled at 1 degree intervals, on a hexagonal grid, before being fed into a 2-dimensional motion detector array with subunits organized as in Fig. 1B. Fig. 3 shows the output of this basic model, plotted as a function of the speed of the pattern and averaged across all parts of the scene. As expected from a system that utilizes multiplication of the inputs, the large range of contrast in the original images leads to an exponential increase in the range of EMD model output responses. At any given velocity, the output of the EMD array is thus highly ambiguous.

**TABLE I**

<table>
<thead>
<tr>
<th>Image</th>
<th>Mean Luminance (Cd/m²)</th>
<th>Global Contrast (SD/mean)</th>
<th>Row Contrast (Mean horizontal 1D power)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A) Barr-Smith</td>
<td>9250</td>
<td>0.807</td>
<td>1.15</td>
</tr>
<tr>
<td>B) Block</td>
<td>28100</td>
<td>0.748</td>
<td>0.783</td>
</tr>
<tr>
<td>C) Botanic</td>
<td>357</td>
<td>3.617</td>
<td>2.90</td>
</tr>
<tr>
<td>D) Bushes</td>
<td>877</td>
<td>3.047</td>
<td>1.671</td>
</tr>
<tr>
<td>E) Car Park</td>
<td>276</td>
<td>4.41</td>
<td>2.13</td>
</tr>
<tr>
<td>F) Close</td>
<td>1690</td>
<td>1.60</td>
<td>1.83</td>
</tr>
<tr>
<td>G) Creek Bed</td>
<td>490</td>
<td>4.642</td>
<td>2.656</td>
</tr>
<tr>
<td>H) Lab</td>
<td>492</td>
<td>2.46</td>
<td>2.05</td>
</tr>
<tr>
<td>I) Mt Lofty</td>
<td>3340</td>
<td>0.932</td>
<td>1.39</td>
</tr>
<tr>
<td>J) Outdoor</td>
<td>5110</td>
<td>0.889</td>
<td>1.24</td>
</tr>
<tr>
<td>K) Rock Garden</td>
<td>3900</td>
<td>1.14</td>
<td>1.45</td>
</tr>
<tr>
<td>L) Shadow</td>
<td>2720</td>
<td>1.45</td>
<td>1.67</td>
</tr>
<tr>
<td>M) Tree</td>
<td>1120</td>
<td>3.32</td>
<td>2.16</td>
</tr>
</tbody>
</table>

**PHOTORECEPTOR+ EMD MODEL**

To test the effect of adding the biomimetic model for the non-linear luminance coding of the insect photoreceptors we then repeated the simulations using a model that incorporated the photoreceptor stage illustrated in Fig 1A. The results are shown in Fig. 4. It can be seen that this dramatically reduced the variability of the responses compared with Fig. 3, and hence with the predictions of the Correlation model. The range of output velocities reduced from 52 dB (40 dB without Block) to 25 dB (8.8 dB without Block).

**BIOLOGICAL DATA**

We recorded intracellularly from an HSNE neuron in the hoverfly *Eristalis tenax* whilst displaying the same set of natural image panoramas to the fly. The HSNE neuron is an identified lobula plate tangential cell (LPTC) with a dorsal receptive field, which stretches along the equatorial axis [17]. This neuron shows responses that are matched to rotational self-motion and is thus tuned to motion along the same axis as our simulations (horizontal). Membrane potentials, normalized relative to rest are shown in Fig. 5. The responses show steady-state examples to simulated yaw rotation (horizontal motion) to the image set from an individual HSNE neuron. Details of the experimental preparation and stimulus display are as described in Reference [18].
The biological data highlights the remarkable consistency in the response to images that otherwise would be expected to confound velocity analysis using ‘basic’ correlation EMDs, with a response range of only 4.6 dB (2.4 dB without ‘Block’). Although in general relatively invariant to different scenes, the biological data revealed the Block image (Fig. 2B) to be a clear outlier. This image was collected specifically to explore the limits of response across different images [10] to a similar level as would be expected for an unelaborated correlation based system (e.g. Fig. 3).

FURTHER MODEL ELABORATIONS

The model motion detection can be further elaborated by the inclusion of a number of steps known or suspected to occur in the insect vision pathway. Such steps included the high-pass spatial and temporal filtering introduced by the second order interneurons, the lamina monopolar cells, as well as saturation stages and motion adaptation. These adaptations are all probable to further improve the robustness of the response across different images [10] to a similar level of the biological system (Fig. 5).

ACKNOWLEDGMENTS

Funding was received from the Australian Research Council (PhD Scholarship to PDB and Post Doctoral Research Fellowship to RSAB) and the US Air Force Office of Scientific Research (FA 9550–04–1–0294).

REFERENCES


