TEMPORALLY FAITHFUL REPRESENTATION OF SALIENT STIMULUS MOVEMENT PATTERNS IN THE EARLY VISUAL SYSTEM

Andreas Thiel[1], Stefan D. Wilke[1],
Martin Greschner[2], Markus Bongard[2],
Josef Ammermüller[2], Christian W. Eurich[1],
and Helmut Schwegler[1]

[1]Institut für Theoretische Physik,
Universität Bremen,
Postfach 330 440, D-28334 Bremen, Germany.
{athiel, swilke, eurich, schwegler}@physik.uni-bremen.de
[2]AG Neurobiologie, FB7,
Carl von Ossietzky Universität Oldenburg,
Postfach 2503, D-26111 Oldenburg, Germany.
martin.greschner@informatik.uni-oldenburg.de,
{markus.bongard, josef.ammermueller}@uni-oldenburg.de

INTRODUCTION

In human visual search, the time required to detect an object that starts to move is approximately independent of the number of distractors in the visual field (pop-out effect)[1]. Motion onset is therefore considered as being a "basic feature". Several theoretical models of the visual system attempt to explain the interplay of such basic features, attentional selection, and higher-level processing[2,3,4,5]. Local contrast of basic features is thought to be calculated in parallel across the visual field in a first step. On the basis of this feature contrast computation, salient positions are determined. The most salient one is then selected by some form of winner-take-all mechanism, the output of which is used to direct attention to this potentially most interesting part of the visual scene.

As salient visual stimuli are first processed by the retina, we use multi-electrode recordings to determine the response of turtle retinal ganglion cells (RGCs) to bright bars rapidly starting and stopping their motion. We describe the measured population activity with a computational model that includes a nonlinear negative feedback loop

originally used for contrast gain control (CGC). As the retinal activity itself does not exhibit the characteristics needed to signal motion onset saliency, we finally apply two adaptive mechanisms by which the retinal response is tranformed into a feature contrast signal. This may be used within the framework of the attentional models described above to guide attention towards objects suddenly starting to move.

RETINAL RESPONSE TO STEPWISE MOTION

Experimental paradigm

Experiments were performed on isolated turtle retinae. The instantaneous discharge rate of a population of ganglion cells was computed after recording their action potentials with a multi-electrode array. Details of the experimental preparation, the recording setup, and the data processing are described elsewhere[6]. The stimulus used was a bright bar (width = 0.1 mm) on a dark background projected onto the retina. This bar was moved uniformly with a speed of 0.44 µm/ms in the direction perpendicular to its orientation. Motion lasted for an interval of 500 ms, after which the bar remained stationary for another 500 ms. Following this pause, motion continued for 500 ms until the next pause. This cycle was repeated until the stimulus had crossed the whole retina.

Computational retina model

A computational model is used to describe the observed retinal population firing rate. The RGC activation $u(\mathbf{r},t)$ at position $\mathbf{r}=(x,y)$ (in µm on the retina) at time t is obtained by convolving the spatiotemporal contrast pattern of the stimulus $s(\mathbf{r},t)$ with a kernel function $K(\mathbf{r},t)$:

$$u(\mathbf{r},t) = g(\mathbf{r},t)\,[K(\mathbf{r},t) * s(\mathbf{r},t-\delta t)]\ . \tag{1}$$

$g(\mathbf{r},t)$ is a modulation factor which will be specified below, $\delta t > 0$ is the response latency, and "$*$" denotes the convolution operator.

$K(\mathbf{r},t)$ is designed to mimic the RGCs' receptive field (RF) properties. Assuming space-time separability, the kernel is split into a spatial part described by a difference of Gaussians and a temporal component with high-pass characteristics[7]:

$$K(\mathbf{r},t) = K_{\mathrm{s}}(\mathbf{r})\,K_{\mathrm{t}}(t)\,, \tag{2}$$

$$K_{\mathrm{s}}(\mathbf{r}) = \frac{g_+}{2\pi\sigma_+^2}\exp\left(-\frac{\mathbf{r}^2}{2\sigma_+^2}\right) - \frac{g_-}{2\pi\beta^2\sigma_+^2}\exp\left(-\frac{\mathbf{r}^2}{2\beta^2\sigma_+^2}\right)\,, \tag{3}$$

$$K_{\mathrm{t}}(t) = \delta(t) - \alpha H(t)\exp(-\alpha t)\,. \tag{4}$$

$H(\cdot)$ denotes the Heaviside step function. The parameters g_+ and g_- determine the relative weights of center and surround, respectively, while σ_+ and $\beta\sigma_+$ (with $\beta > 1$) are their diameters. In the temporal part, α^{-1} is the decay time constant of the response.

From the activation $u(\mathbf{r},t)$, the instantaneous population firing rate is computed by a rectification to exclude negative firing rates, and a spatial integration over the retinal area A considered.

$$f(\mathbf{r},t) = \tilde{\alpha}\ [u(\mathbf{r},t)+\Theta]_+\ , \tag{5}$$

$$R(t) = \int_A f(\mathbf{r},t)\mathrm{d}\mathbf{r}\,, \tag{6}$$

where $\tilde{\alpha}$ and $\Theta \geq 0$ determine scale and baseline of the firing rate, respectively; $[x]_+ := xH(x)$ is the rectification operator. Apart from $\tilde{\alpha}$, Θ and α, the numerical values of the parameters are not critical for the model's temporal characteristics. Thus, they are matched to results from other experiments: $\beta = 3$ and $g_- = 0.8\, g_+$ is chosen for the spatial part of the kernel[7]. This is consistent with the study of Smirnakis et al.[8], yielding a reasonable excitatory/inhibitory RF weight ratio of $s = 1.55$. A RF radius of $\sigma_+ = 80\,\mu\text{m}$ is chosen[8].

A CGC feedback loop is included in the model via the modulation factor $g(\mathbf{r}, t)$ in equation 1. $g(\mathbf{r}, t)$ is computed by first feeding the activity $u(\mathbf{r}, t)$ into a temporal low-pass filter and then applying a static nonlinearity to the filtered result:

$$v(\mathbf{r}, t) := B\, u(\mathbf{r}, t) * [H(t) \exp(-t/\tau)]\,, \tag{7}$$

$$g(\mathbf{r}, t) = \frac{1}{1 + \{[v(\mathbf{r}, t)]_+\}^4}\,. \tag{8}$$

The parameters B and τ govern the strength and the time course of the CGC modulation, respectively. Functionally, this feedback loop generates a delayed local suppression of high, sustained RGC activation, thus altering the temporal characteristics of RGC firing rates. The form of CGC described above was originally suggested by Berry II et al.[9] and used to explain retinal motion anticipation effects. Setting the parameters $g_+ = 3$, $\alpha = 4\,\text{Hz}$, and $\tau = 170\,\text{ms}$ reproduces the qualitative behavior shown in Figure 3 of the abovementioned publication[9]. The low decay rate $\alpha = 4\,\text{Hz}$ is crucial for our conclusions, we therefore verified it by comparing it to values obtained from flash stimulation[10] and reverse correlation[8].

Experimental results and parameter adjustment

The mean population discharge rate in response to a succession of stepwise uniform movements is shown in Figure 1. After the onset of motion, there is a response latency of about 100 ms. Within another 100 ms, the retina signals motion onset by a 100% rise in activity. During the time following this signal, the population activity remains fairly constant, corresponding to the response to continuous movement. At motion offset, the retinal activity exhibits an equally steep drop and thus closely follows the movement pattern.

We reproduce the observed behavior with the retina model by adjusting its parameters B, Θ, and $\tilde{\alpha}$. Parameter fitting is performed by minimizing the average squared difference between experimental data and model prediction according to three scenarios: Scenario I represents the full model with CGC. To demonstrate the CGC's ability of accelerating the response, it was "turned off" in scenario II by setting $B = 0$. Finally, scenario III allows the linear time constant α to be changed as well while still $B = 0$.

Comparing the ability of the three model scenarios to describe the time course of retinal population activity, one finds that only scenario I yields good agreement between model and experimental data (adjusted parameters: $\tilde{\alpha} = 79\,\text{Hz}$, $B = 78\,\text{Hz}$, $\Theta = 0.005$). Scenario II fails to exhibit a steep slope at beginning motion ($\tilde{\alpha} = 30\,\text{Hz}$, $\Theta = 0.015$). Although the unusually large value of α in scenario III ($\tilde{\alpha} = 104\,\text{Hz}$, $\Theta = 0.005$, $\alpha = 24\,\text{Hz}$) results in a sufficiently quick rise at motion onset, activity decays much slower after motion stops in reality than predicted by scenario III. A point not captured by any of the models is that the maximum signal is generated *early* in each motion phase. Instead, the models yield the highest response at the end of the movement.

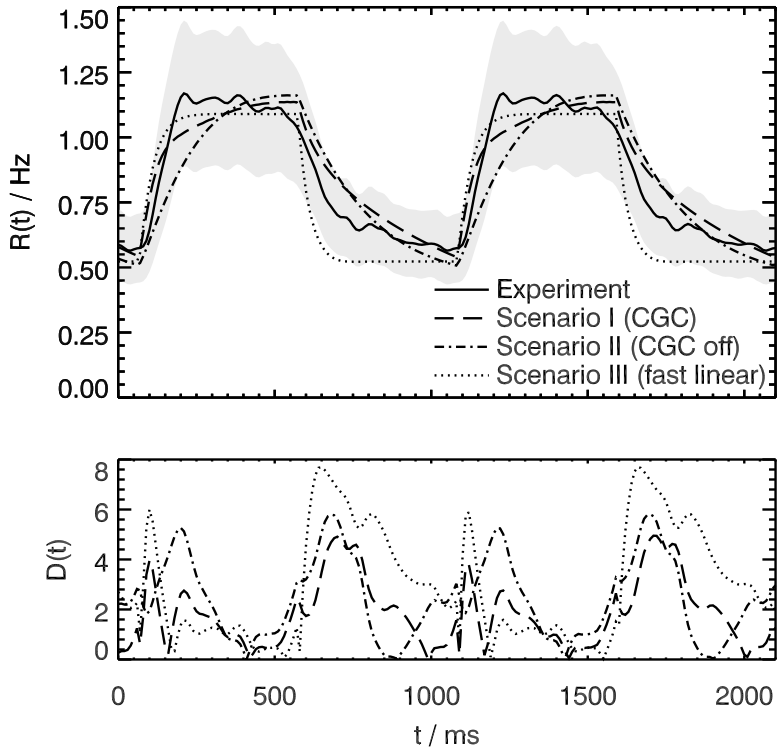

Figure 1: **Upper panel:** Population response of turtle retinal ganglion cells (solid, 219 RGCs, 450 stimulus repetitions) to stepwise motion of a bright bar (width 100 µm, periods of movement from $t = 0$ ms to $t = 500$ ms and from $t = 1000$ ms to $t = 1500$ ms) as a function of time. Shaded area indicates standard error. Corresponding output of the retina model with contrast gain control loop (scenario I, dashed), without contrast gain control (scenario II, dot-dashed), and without contrast gain control but with adjusted linear time constant (scenario III, dotted). **Lower panel:** Deviation $D_i(t)$ of the mean experimental population discharge rate $R_{\mathrm{exp}}(t)$ and scenario i's prediction $R_i(t)$: $D_i(t) := \sqrt{(R_{\mathrm{exp}}(t) - R_i(t))^2}/\sigma_{\mathrm{m}}(t)$ for $i = $I,II,III, where $\sigma_{\mathrm{m}}(t)$ denotes the standard error of $R_{\mathrm{exp}}(t)$.

SALIENCY COMPUTATION

Differentiation by adaptive mechanisms

An assumption used in most attentional models is that saliency is coded in the firing rate of the neurons forming the feature map. From this viewpoint, the population of RGCs itself does not generate a signal that expresses the saliency of motion onsets, as activity right after the bar has started to move is only slightly higher than during continuous stimulation (see Figure 1). Nevertheless, the faithful representation of movement by the retinal activity can be used as a basis for saliency computation. Mathematically, the signal leaving the retina must be differentiated to detect motion onsets. This differentiation can approximately be accomplished by neurons in the next processing stage that exhibit fatigue or depletion when receiving constantly high retinal input but recover during low activity. Such "postretinal neurons" (PRNs) should

therefore be sensitive to the steeply rising activity at motion starts because these are preceded by pauses due to motion stops. In particular, depressive synaptic transmission and fatigue in spike generation are plausible candidate mechnisms with the characteristics mentioned above. PRNs that are sensitive to temporal structure of their input in this way are capable of computing motion onset saliency. They may then be arranged in a retinotopic feature map to represent the position of the salient stimulus, to select the most salient position, and to direct attention to that position in the same way used in former attentional models. Neurons of this kind may be found in the superior colliculus, where they could direct saccades to objects suddenly starting to move.

Depressive synaptic transmission

Depressive synaptic transmission is modeled according to Tsodyks and Markram[11]. The mean postsynaptic current (PSC) $P(\mathbf{r},t)$ of a single synapse in response to the RGC firing rate $f(\mathbf{r},t)$ is proportional to

$$P(\mathbf{r},t) \propto \tau_{\text{in}} U f(\mathbf{r},t) \frac{1}{\tau_{\text{rec}}} \int_{-\infty}^{t} \mathrm{d}t' \exp\left[-\frac{t-t'}{\tau_{\text{rec}}} - U \int_{t'}^{t} \mathrm{d}\theta \; f(\mathbf{r},\theta) \right] , \tag{9}$$

Upon the arrival of each presynaptic spike, a certain amount of neurotransmitter U (the "use parameter") is released, which is replaced afterwards with constant rate $1/\tau_{\text{rec}}$. If spikes arrive too often, depletion dominates regeneration and prevents the transmission of the presynaptic signals. In contrast, rapid increases in firing rate preceded by low activity lead to high PSCs because the amount of available transmitter has had the chance to recover[12]. Figure 2 shows the PSC calculated from the model responding to a single motion onset for different use parameter values. In the simulations, we chose the

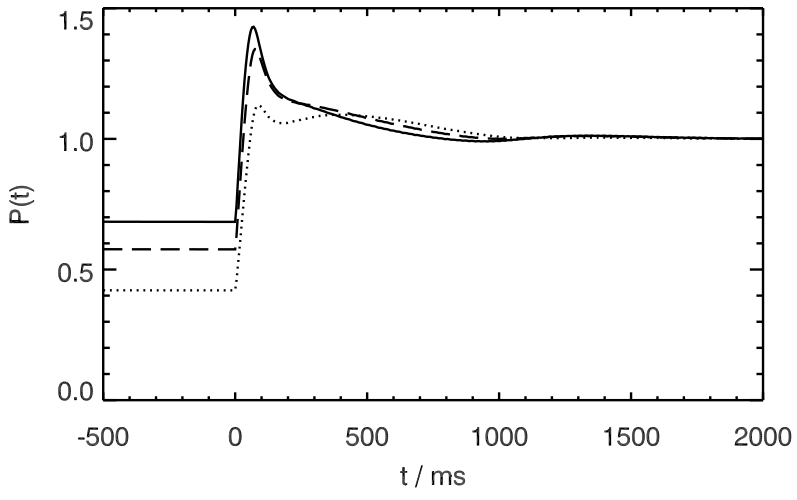

Figure 2: Normalized mean postsynaptic current (PSC) of depressive synapses evoked by motion onset responses of simulated retinal ganglion cells as a function of time. The mean PSC in response to continuous motion was normalized to unity. Curves were obtained with use parameter $U = 0.1$ (dotted), $U = 0.5$ (dashed), and $U = 0.9$ (solid).

generic values $\tau_{\text{in}} = 3\,\text{ms}$ and $\tau_{\text{rec}} = 800\,\text{ms}$ given by Tsodyks and Markram[11], whereas U varied between 0.1 and 0.9. The PSCs at all mesh points were averaged at each time step to obtain a measure of mean population PSC. The low activity before motion starts allows for transmitter recovery, such that immediately after motion start the PSC exhibits an overshoot that becomes more prominent with increasing use parameter.

Adaptive action potential generation

To analyze processing of motion information in PRNs, we simulated a population of leaky integrate-and-fire neurons ($N = 20$) with adaptive spike generation using a dynamic threshold[13,14]. Single cell spiking behavior and population activity in response to light bars crossing the cells' RFs are compared for continuous as opposed to stepwise stimulus motion. Assuming that one PRN receives signals from a whole population of RGCs due to dendritic integration, the measured retinal activity can be used as the input signal $I(t)$ to one of the model cells. Cell i's membrane potential $M_i(t)$ is then calculated by low-pass filtering this input signal,

$$M_i(t) = I(t) * \left[H(t)\exp(-t/\tau_m)\right]. \tag{10}$$

When $M_i(t)$ exceeds cell i's threshold $\theta_i(t)$, a spike is generated in the following simulation time step $t + \Delta t$:

$$O_i(t + \Delta t) = H(M_i(t) - \theta_i(t)). \tag{11}$$

$\theta_i(t)$ is the sum of a static offset θ_0 and two dynamic parts θ_r and θ_s. Both of them are increased each time cell i generates an action potential and decay exponentially afterwards:

$$\theta_i(t) = \theta_0 + O_i(t) * \left[H(t)(V_{\theta r}\exp(-t/\tau_{\theta r}) + V_{\theta s}\exp(-t/\tau_{\theta s}))\right]. \tag{12}$$

θ_r has a large increase factor $V_{\theta r} = 1.0$ and a short time constant $\tau_{\theta r} = 20\,\text{ms}$, modeling the cell's relative refractory period. θ_s's increase factor $V_{\theta s} = 0.1$ is smaller, while its time constant $\tau_{\theta s} = 300\,\text{ms}$ is larger, mimicking slowly accumulating fatigue in spike generation.

In the stepwise motion column of Figure 3, times (1) and (2) indicate the start and stop of a single motion step respectively. During this time interval, cell $i = 10$ fires several spikes. Although in the case of continuous motion, the bar passes the same receptive field positions with the same velocity between the corresponding times (1) and (2), only one action potential is generated. The single cell's behavior is reproduced in the activity of the population of PRNs. Population activity is much higher during stepwise stimulation compared to the continuous condition.

The model cells' behavior is caused by the recovery of the spike generation mechanism. In the case of continuous motion, the constant retinal input results in a high membrane potential, repetitive firing, and an increase of the slow threshold component θ_s, thereby reducing the number of spikes generated. The low retinal input during motion pauses allows θ_s to decrease, such that immediately after the following motion onset, spike generation is facilitated until θ_s has built up again.

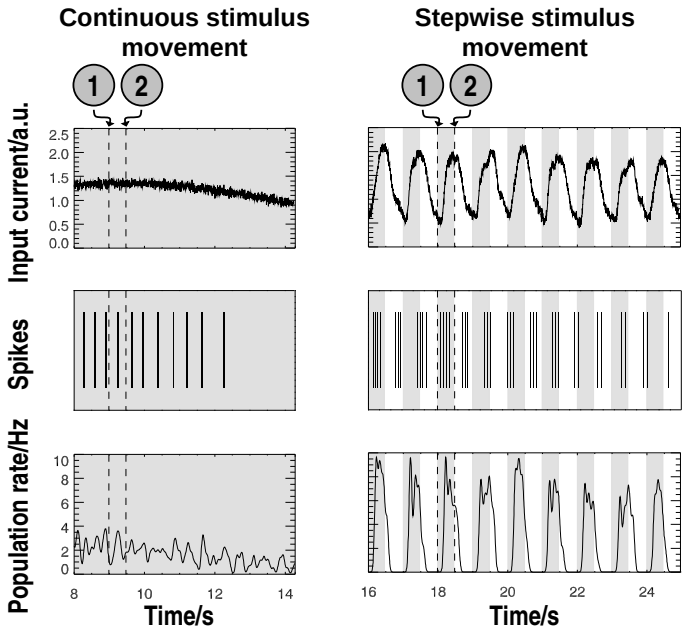

Figure 3: Input current (top) and action potentials generated by a single (neuron $i = 10$, middle) and by a population ($N = 20$, bottom) of postretinal model cells responding to two different motion patterns (left and right column). In the stepwise motion column, motion periods are depicted by gray areas, and times (1) and (2) indicate the start and stop of a single motion step respectively. In the case of continuous motion, the bar passes the same receptive field positions with the same velocity at corresponding times (1) and (2), but without preceding and subsequent pausing.

SUMMARY

By recording ganglion cell population discharge rates using a multi-electrode array, we have shown that the time course of this activity closely reflects rapid velocity changes of a stepwise moving stimulus. An extremely sharp rise in the activity signals motion onsets, but its amplitude does not increase beyond the level measured in response to continuous movement. The steep and quick rise of the RGC population rate was described using a computational model with a CGC feedback loop. Models without this non-linear element can also explain the fast response, but at the expense of very small time constants, which were shown to be inconsistent with our and several other previous studies. Thus, CGC plays a crucial role in the detection of movement onset.

Moreover, we have shown that the RGC response to motion patterns can be transformed into a saliency signal with sharp peaks at each motion onset by neurons in the following processing stage. This was demonstrated using two different mechanisms: Depressive synaptic transmission and dynamic thresholds in PRNs. In both cases, an adaptive element inhibits the high-amplitude retinal input during the movement periods but recovers during the pauses, which allows for a strong response upon the subsequent motion onset. Neurons sensitive to temporal feature contrast in this way could then be the components of a retinotopic feature map. An attentional model including this feature map would exhibit a preference for motion onsets and could thus be capable of explaining the high saliency of objects that suddenly start moving.

REFERENCES

1. J. M. Wolfe. Visual search. In H. Pashler, editor, *Attention*, pages 13–74. Psychology Press, 1998.

2. C. Koch and S. Ullman. Shifts in selective visual attention: Towards the underlying neural circuitry. *Human Neurobiology*, 4:219–227, 1985.

3. E. Niebur, C. Koch, and C. Rosin. An oscillation-based model for the neuronal basis of attention. *Vision Research*, 33(18):2789–2802, 1993.

4. B. A. Olshausen, C. H. Anderson, and D. C. van Essen. A neurobiological model of visual attention and invariant pattern recognition based on dynamic routing of information flow. *Journal of Neuroscience*, 13:4700–4719, 1993.

5. R. Goebel. *Visuelle Aufmerksamkeit, perzeptive Organisation und invariante Objekterkennung: Eine Integration neurobiologischer und psychologischer Befunde in einem neuronalen Netzwerk-Modell*. PhD thesis, TU Braunschweig, 1995.

6. S. D. Wilke, A. Thiel, C. W. Eurich, M. Greschner, M. Bongard, J. Ammermüller, and H. Schwegler. Population coding of motion patterns in the early visual system. Submitted to *Journal of Computational Neuroscience*, 2000.

7. R. W. Rodieck. Quantitative analysis of cat retinal ganglion cell response to visual stimuli. *Vision Research*, 5:583–601, 1965.

8. S. M. Smirnakis, M. J. Berry, D. K. Warland, W. Bialek, and M. Meister. Adaptation of retinal processing to image contrast and spatial scale. *Nature*, 386:69–73, 1997.

9. M. J. Berry II, I. H. Brivanlou, T. A. Jordan, and M. Meister. Anticipation of moving stimuli by the retina. *Nature*, 398:334–338, 1999.

10. A. L. Jacobs and F. S. Werblin. Spatiotemporal patterns at the retinal output. *Journal of Neurophysiology*, 80:447–451, 1998.

11. M. V. Tsodyks and H. Markram. The neural code between neocortical pyramidal neurons depends on neurotransmitter release probability. *Proceedings of the National Academy of Sciences USA*, 94:719–723, 1997.

12. M. Bethge, K. R. Pawelzik, and T. Geisel. Brief pauses as signals for depressing synapses. *Neurocomputing*, 26-27:1–7, 1999.

13. R. Eckhorn, H. J. Reitboeck, M. Arndt, and P. Dicke. Feature linking via synchronization among distributed assemblies: Simulations of results from cat visual cortex. *Neural Computation*, 2:293–307, 1990.

14. M. J. M. Lankheet, J. Molenaar, and W. A. van de Grind. The spike generating mechanism of cat retinal ganglion cells. *Vision Research*, 29:505–517, 1989.