A Computational Framework for a Bio-inspired Mechanism of Vernier Hyperacuity

András Róka, Ákos Zarányi

Cellular Sensory and Wave Computing Laboratory, Computer and Automation Research Institute of the Hungarian Academy of Sciences Budapest University of Technology and Economics, Hungary
e-mail: andras.roka@sztaki.hu, zarandy@sztaki.hu

Abstract: In this paper, we present a bio-inspired computational framework for the hypothetical cooperation of the ocular drift and a place coding neural circuitry. The proposed bio-inspired mechanism can provide an explanation for the ability of primate vernier hyperacuity. Our starting point is a spatiotemporal model of the primate retinal P and M ganglion cells. First we show that in the central retina, the drift-induced movement moves the stimuli of preferred spatial frequency with a velocity that is optimal for P cells and almost "invisible" for the movement-sensitive M cells. Secondly, based on a presumed analogy between the owl's auditory system and a primate visual system we present a theory that visual hyperacuity can be originated to the detection of time delays between neural firing patterns. Third, we propose a possible neural model for the place coding circuitry. Our new theory may serve suggestions for further research.

Keywords: ocular drift; place coding circuitry; hyperacuity; computational model

1 Introduction

Vernier acuity is a measure of the smallest offset between two line segments that can be discriminated. Hyperacuity is the ability of the eye to resolve an offset with a resolution better than that imposed by the Nyquist limit of the photoreceptor mosaic. Humans can resolve details with an accuracy of better than one fifth of the size of the most sensitive photoreceptor. The explanation for this remarkable ability has been under debate from its discovery.

At present, a wide range of computational models exists for visual hyperacuity and its perceptual learning mechanism. Without the aim of completeness, we survey the recent relevant studies. One of these states that saccadic eye movements may improve the resolution of the eye. During fixation, the eyes "sample" a scene through microsaccades. These "sampled" images can provide the human brain enough information to resolve hyperacuity tasks [35]. Another theory
assumes that the spatial fluctuations of retinal sampling and the temporal fluctuations caused by eye-tremors can induce noise-enhanced processing effects, providing the hyperacuity phenomenon [71]. Furthermore, a model in [70] states that hyperacuity and its dependence on stimulus length can be realized by the integration of information from more than a single neuron.

In this work we present a bio-inspired theory about the perceptual mechanism of primate vernier hyperacuity. We assume that the small spatial offset in vernier acuity tasks is converted into an appropriate time offset between the firing patterns of the corresponding neurons. We show that an involuntary eye-movement – the ocular drift – may be responsible for this conversion. The proposed mechanism and the underlying neural circuitry makes it possible to convert unresolvable spatial offsets into resolvable time differences. As far as we know, there are no similar theories that would find a functional connection between ocular drift and visual hyperacuity.

The paper is structured as follows. In Section 2 we give a brief overview of our model retina. The model is comprehensible without a detailed knowledge of retinal functionality. This section ensures only the reproducibility of our numerical simulations. Sections 3 and 4 present an overview of the properties of ocular drift and visual hyperacuity. In Section 5 we introduce our bio-inspired computational model and show our simulation results. Finally the occurring questions are discussed in Section 6.

2 Methods

In this study we use a computational model of the light-adapted primate retina regarding to the four main type of retinal ganglion cells (the ON and OFF subtypes of P and M ganglion cells). The model aims to establish a comprehensive accumulation and representation of knowledge about many aspects of the primate retina, including details that are directly reported in the literature, as well as those that are carefully inferred from studies in other species. The model converts the retinal morphology into a transfer function representation. Here we give a brief description of our model retina.

1) Our starting point was the analytical approximation of cone and total ganglion cell density along the surface of the entire human retina. Both approximations are based on the measured dataset of Curcio et al. [12, 13]. Results were compared with others [69, 25, 2, 45]. Regarding to ganglion cells, the method of Drasdo et al. [19] was used to handle the effect of lateral displacement. The resulted foveal ganglion-cell-to-cone ratio was compared with measured values [13, 54, 55, 19, 42, 50, 60, 63, 29, 1].
2) We calculated the ratio of P, M and K cells\(^1\). The formulas are based on the estimations in [50, 15, 3] and have been validated by the cumulative number and ratio of the RGC types [26, 14, 15, 41].

3) The ratio of ON and OFF cells to the total P and M cell population were also estimated corresponding to the observations in the literature [15, 16, 1, 33, 19].

4) From the density of the different cell types, the spacing of the receptive field mosaic was determined in function with eccentricity. Using the measured angle of RF\(^2\) center's overlap [27, 28, 8, 65], the radii of the RF centers can also be estimated. Given the radius of each cell's RF center and considering the experimentally measured center/surround ratio [11], we computed the full extent of the receptive fields.

5) Using the well-known Difference-Of-Gaussians modeling approach [47], the spatial sensitivity functions of P and M cells can also be computed [64, 11, 22].

6) Regarding the temporal sensitivity functions of these ganglion cells, we used well-known measurements and approximations. In [32, 6, 7] the authors give a transfer function representation for P and M cell RF center and surround. The measurements were made on Macaca fascicularis, using drifting sinusoidal gratings at 1123 td (P cells) and 1180 td (M cells) retinal luminance levels.

7) The general form of estimating RGC response is as follows [38, 31].

\[
R = (S_{rfc}(v)T_{rfc}(j\omega) - S_{rfs}(v)T_{rfs}(j\omega))LC
\]

Where \(R\) is the response [imps/sec] of a ganglion cell, \(S_{rfc}\) and \(S_{rfs}\) are the spatial sensitivities, \(T_{rfc}\) and \(T_{rfs}\) are the temporal sensitivities of the RF center and surround respectively. \(L\) and \(C\) are stimulus properties: the mean retinal illuminance and the Michelson contrast. Spatial and temporal frequencies are denoted by \(v\) and \(\omega\).

Figure 1 shows the spatial sensitivity functions of two types of retinal ganglion cells (OFF-P and OFF-M), calculated from our model. The boundaries of the colored domains represent the spatial frequencies corresponding to the -3dB and -6dB values of the spatial sensitivity functions. The density of both P and M cells rapidly decreases toward the periphery, with the result that the size of the receptive fields increases and thus the preferred spatial frequencies decreases significantly. The curves in Figure 1 well correlate with the findings that human high-contrast acuity is about 60 cycles/degree (P cells at zero eccentricity) [19], while low contrast sensitivity has a maximum around 3 cycles/degree at the fovea.

---

RGC: Retinal Ganglion Cell

\(^2\) RF: Receptive Field
(M cells at zero eccentricity) and decreases to about 0.2-0.3 cycles/degree at the far periphery [4].

Figure 1
Spatial sensitivity function of two RGC types (OFF-P and OFF-M) in function with eccentricity. The boundaries of the red and yellow domains represent the spatial frequencies corresponding to the -3dB and -6dB values of the spatial sensitivity.

Figure 2
Left: A typical linear filter model to P cell response in the frequency domain. Right: A typical linear filter model to M cell response at different contrast levels (0.05-0.4). The contrast adaptation modifies the dynamics of the M cells. With increasing local contrast, the M cell peak temporal frequency shifts from ~9 Hz to ~17 Hz. The filled domain represents the preferred (-3dB) frequency band of the cells.

Stimuli: drifting sinusoidal grating, spatial frequency was optimal for the cells. Subject: Macaca fascicularis. Retinal luminance level: 1123 td (left) and 1180 td (right).
Figure 2 shows the typical temporal sensitivity functions of P and M cells for achromatic stimuli at photopic light levels. In the primate retina M cells show contrast gain control [21, 53, 31, 6], while P cells do not [38, 53, 32]. The contrast gain control modifies the dynamics of the M cells. With increasing local contrast, the M cell peak temporal frequency shifts from $\sim 9 \text{ Hz}$ to $\sim 17 \text{ Hz}$ [17, 5, 38, 7].

Given the model retina, we can calculate the effect of involuntary eye movements on the retinal cells. In the following section we give a brief description of the ocular drift and we analyze the effect of drift-induced stimulus movement on the P and M ganglion cells.

3 Ocular Drift

Drifts are slow involuntary motions of the eye occurring between microsaccades, simultaneously with tremor. During drifts, the image of the object being fixated can move across several photoreceptors [46] (Figure 3). Initially, drifts seemed to be random motions of the eye generated by the instability of the oculomotor system [18, 10], but later were found to have a compensatory role in maintaining accurate visual fixation in the absence of microsaccades, or at times when compensation by microsaccades was relatively poor [57].

The mean amplitude of the ocular drifts is between 1.2 and 6 arcmin, while the average velocity ranges between 6 and 12 arcmin/sec [39]. It is worth noting that – like a "rule" – the drift reaches its almost maximal velocity (close to 30 arcmin/sec), several times per second [68].

![Figure 3](image)

Involuntary eye-movements: Slow drifts (curved lines) and tremor (superimposed on drifts). Drift periods are interrupted by microsaccades, the fast jerk-like eye-movements.
Using our retina model, we can calculate the effect of the ocular drift on P and M ganglion cell response. The motion of the eye evidently induces a motion of the stimulus on the retinal surface. An achromatic contrast stimulus – traveling on the retina with the drift velocity – represents a dynamic input for a retinal cell. Without loss of generality, assuming a sinusoidal (spatial) contrast grid, a ganglion cell receptive field receives a time-varying sinusoidal contrast input. The temporal frequency of the contrast input can be calculated by multiplying the grid spatial frequency with the drift velocity.

We compared the ganglion cells’ preferred temporal frequency domain with the “drift-induced” temporal frequency. In Figure 4 the thick black lines represent the cell’s preferred spatial frequency multiplied by the average drift velocity. The narrow domains show the -3dB spatial frequency band multiplied by the average drift velocity, while the wider domains represent the same spatial frequency band multiplied by the upper and lower limits of drift velocity. The gray domains show the preferred temporal frequency bands of M and P cells respectively (as shown in Figure 2).

![Figure 4](image)

**Figure 4**
Temporal frequencies induced by ocular drift for M (left) and P (right) cells. The gray domain shows the cells’ preferred frequency bands as shown in Figure 2.

In the central retina, M cells have much larger receptive fields compared to P cells. This property makes the M cells sensitive to lower spatial frequencies (~3 cycles/degree). In Figure 4 it clearly seems that the relatively slow ocular drift cannot move the stimuli of these low spatial frequencies fast enough to produce temporal frequencies in the cell’s preferred frequency domain. In other words, M cells cannot respond vigorously to the drift-induced motions of the preferred stimuli. After all it is that a cell type involved in movement detection would not be “misled” by an involuntary eye movement.

P cells – which have a very small receptive field – are sensitive to high spatial frequencies (up to 60 cycles/degree). These fine details – moving by the drift velocity – produce temporal frequencies high enough to elicit considerable cell responses. It means that the drift induced motion may be appropriate for producing P cell activity (Figure 4).
The knowledge of the spatial and temporal properties of retinal cells gives us the opportunity to calculate the human visual acuity, but the phenomenon of hyperacuity cannot be explained based on these data. The next section gives a short overview of human visual hyperacuity and the neural basis of the owl's auditory hyperacuity.

4 Hyperacuity

The term "sensory acuity" refers to the ability of the brain to resolve fine details. Visual acuity is the ability to discriminate the finest detail, for example two parallel lines apart. In humans it is limited by the sharpness of the retinal focus and the number of photoreceptors together. The human foveal visual system's acuity is about 1 min of arc (60 cycles/degree).

Hyperacuity is the ability of sensory modalities to detect differences in two or more stimuli well below the sensory resolution. Humans can resolve details with an accuracy of better than one fifth of the size of the most sensitive photoreceptor. Figure 5 shows a typical hyperacuity task.

![Figure 5](image)

A typical hyperacuity task: reading a sliding caliper

Hyperacuity has been studied for over a hundred years and a wide range of studies provide computational models trying to imitate and explain this phenomenon. However, many aspects of the underlying mechanism are still debated.

The auditory hyperacuity of the barn owl has also been investigated for decades [58, 34]. When localizing its pray, relying solely on acoustic signals, the owl's auditory system can detect shorter delays of time arrivals of sound than the duration of an action potential, which indicates the time arrival. It has been shown that the small time differences of signals from the left and right ears are mapped into a neural place coding circuitry, containing small internal delays and coincidence detector neurons [9] (see Figure 6). The theoretical model – where the place of the neuron with maximum response specifies the corresponding time delay – was first proposed by Jeffres [30], and experimentally verified by [58, 34, 59]. A more detailed model description is presented in [37].
5 Results

In this section we introduce an analogy between the owl's localization system and the human vernier hyperacuity. The sound signal (coming from a location) approaches the two sets of receptor cells in the owl's left and right ears (Figure 6). Similarly, the stimulus – moving by the ocular drift on the central retina – approaches the two sets of photoreceptors signed with "A" and "B" in Figure 7 (b).

In both cases, the time difference in stimulus arrival originates from spatial displacements: the displacement of the two ears of the owl and the displacement of the two line segments (the displacement is signed with $D$ in Figure 7 a). The value of time difference can be calculated by dividing the spatial displacement
with the stimulus traveling velocity. In the owl's case it can be as little as some tens of microseconds. In the human visual system (assuming $12 \text{ arcmin/sec}$ mean drift velocity and $D = 0.1 \text{ arcmin}$ smallest resolvable spatial offset) the time difference is about 8 ms. The firing pattern of the ganglion cells regarding to the “A” and “B” sets are shown in Figure 7.

In a simulation example, we show that the possible cooperation of ocular drift and a place coding circuitry introduced by [30] is capable of explaining the ability of vernier hyperacuity. We created a computational model of a small part of human foveola, where there are no rods and the great majority ($>95\%$) of ganglion cells are P cells. In this small central area every cone is connected to at least two ganglion cells (an off-type and an on-type ganglion cell) and the center of a midget cell receptive field is fed by a single cone [19].

![Figure 8](image)

**Figure 8**

Firing patterns of “set A” and “set B” ganglion cells to the stimulus presented in Figure 7

The stimulus we applied is a classical instance of vernier hyperacuity tasks. The black lines occupy 1 arcmin in width as shown in Figure 7 a. The applied displacements of the two lines are $D = 0.05, 0.1, 0.15 \text{ and } -0.1 \text{ arcmin}$ respectively. The displacement of $0.1 \text{ arcmin}$ is ten times smaller than the resolving power of the photoreceptor mosaic, and is close to the smallest resolvable difference in hyperacuity tasks. The stimulus is blurred according to the foveal point spread function [19] and moves on the surface of the model retina with 12 arcmin/sec assumed drift velocity (see Figure 7 b).

The firing patterns – evoked by the described stimuli – were used as an input of a coincidence detector circuitry (Figure 9 left side). The delay values were appointed by us to allow 0.05 arcmin spatial resolution in space coding. From the viewpoint of this model, it is favorable that P ganglion cells produce a relatively low firing rate, having 7-10 times lower contrast gain compared to M cells [21, 44, 32]. The lower firing rate enables the place coding circuitry to resolve larger time differences without ambiguity.

As shown in Figure 9, the place coding circuit is able to discriminate different degree of displacements of the two lines segments, achieving the ability of hyperacuity. Theoretically, the proposed circuitry is able to resolve smaller
differences than the human hyperacuity limit, but in a real retina the cell mosaic
does not show this high regularity and drift also has its own variability in velocity
and direction. These effects may limit the resolution of the mechanism.

Figure 9
Simulation example. On the left side the simulated place coding network is presented. The top row
shows the ganglion cell responses regarding to the different stimuli. The firing patterns of the place
coding circuitry are presented in columns.

It is important to note that this model is bio-inspired, since it does not provide
evidence for the existence of this mechanism in the primate brain, but only
proposes a biological concept for achieving visual hyperacuity.

6 Discussion

Our model raises several new questions that we are unable to answer at present. It
is evident that a kind of place coding mechanism would strongly depend on the
drift direction and velocity. If drift velocity is too slow, or drift direction is far
from perpendicular to the line segments, the time differences between firing
patterns can exceed the half of the P cell minimum inter-spike interval, resulting
in a false detection. The problem can be solved by repeating the detection process
with a different drift velocity and/or direction.
There is no evidence for a mechanism that would adjust the drift characteristics in hyperacuity tasks, but it is known that drift shows high variability – reaching maximal velocity values several times per a second [68] – and it is at least indirectly influenced by visual factors [40]. Furthermore it is known that human performance improves with practice (perceptual learning) in hyperacuity tasks [61, 24]. If we assume that perceptual learning have some effects on drift direction and/or velocity then it would explain why learning improvements are highly specific to task, stimulus type, orientation, retinal location and eye trained [56, 23, 43], and why improvements in one task diminish after a similar task has been subsequently practiced [56, 52].

Another open question is the structure of the proposed neural circuitry. If human visual hyperacuity is based on a place coding circuitry, then which exact neural network may serve this function and where does it possibly takes place in the visual pathway?

The internal delays in our proposed circuitry fall in the range of milliseconds, which can be ascribed to both long axons and "delay neurons". The slowest unmyelinated axons with $\sim 0.1 \text{ um}$ diameter have about $0.3 \text{ m/s}$ conduction velocity. The place coding circuitry (shown in Figure 9) needs about 3-4 mm length of this type of axon. Without suggesting a possible location in the visual cortex, we note that this axon length is conceivable. Similar lengths were observed in layer 2/3 of striate cortex, where pyramidal cells extend long axons that form clustered projections linking iso-orientation columns [20].

It is important to note that in our model, the resolution of hyperacuity is related to the number of place coding neurons, which is evidently independent of the ganglion cell density, while acuity is directly related to ganglion cell density. It well correlates with the observation that hyperacuity tasks appear to be limited by the extent of striate cortical representation both in the fovea and in the perifoveal field [36, 66, 67], while contrast and resolution tasks seem to be limited by ganglion-cell density [62, 48, 49].

Conclusions

In this work we propose a new theory about human vernier hyperacuity. The theory states that an involuntary eye movement – the ocular drift – is ideal for converting a small spatial offset (which is unresolvable by the photoreceptor mosaic) into a time offset between firing patterns, which is already resolvable by a place coding neural circuitry. Similar circuitry has been found in owl's brain achieving auditory hyperacuity. In our theory, the perceptual learning has a possible role in adjusting drift direction, duration and velocity to the hyperacuity task. As far as we know, this is the first theory that assumes such an important role to the ocular drift in hyperacuity tasks.

In a computational example we show that the existence of the proposed mechanism is plausible. The theory raises several new questions that we are
unable to answer at present, but we hope they may serve suggestions for further research.

Acknowledgement

We would like to say thank you to Prof. András Radványi and Judit Veres for the help during the preparation of this manuscript.

References


