



HAL
open science

Speed-Selectivity in Retinal Ganglion Cells is Sharpened by Broad Spatial Frequency, Naturalistic Stimuli

Cesar Ravello, Laurent U Perrinet, Maria-Jose Escobar, Adrian Palacios

► To cite this version:

Cesar Ravello, Laurent U Perrinet, Maria-Jose Escobar, Adrian Palacios. Speed-Selectivity in Retinal Ganglion Cells is Sharpened by Broad Spatial Frequency, Naturalistic Stimuli. *Scientific Reports*, 2019, 9 (1), 10.1038/s41598-018-36861-8 . hal-02007905

HAL Id: hal-02007905

<https://amu.hal.science/hal-02007905>

Submitted on 18 Dec 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

SCIENTIFIC REPORTS



OPEN

Speed-Selectivity in Retinal Ganglion Cells is Sharpened by Broad Spatial Frequency, Naturalistic Stimuli

César R. Ravello¹, Laurent U. Perrinet², María-José Escobar³ & Adrián G. Palacios¹

Motion detection represents one of the critical tasks of the visual system and has motivated a large body of research. However, it remains unclear precisely why the response of retinal ganglion cells (RGCs) to simple artificial stimuli does not predict their response to complex, naturalistic stimuli. To explore this topic, we use Motion Clouds (MC), which are synthetic textures that preserve properties of natural images and are merely parameterized, in particular by modulating the spatiotemporal spectrum complexity of the stimulus by adjusting the frequency bandwidths. By stimulating the retina of the diurnal rodent, *Octodon degus* with MC we show that the RGCs respond to increasingly complex stimuli by narrowing their adjustment curves in response to movement. At the level of the population, complex stimuli produce a sparser code while preserving movement information; therefore, the stimuli are encoded more efficiently. Interestingly, these properties were observed throughout different populations of RGCs. Thus, our results reveal that the response at the level of RGCs is modulated by the naturalness of the stimulus - in particular for motion - which suggests that the tuning to the statistics of natural images already emerges at the level of the retina.

Motion detection is essential for animal survival, and studies show that many retinal cells^{1–3} and a considerable portion of the cortex⁴ are involved in its processing. However, most of these studies have focused on the detection of artificial stimuli, such as random dots, lines, or moving gratings, while ignoring the natural signals from the environment in which animals have evolved and to which their visual system would be attuned to^{5–9}. Moreover, computational models based on the response to simple, artificial stimuli fail to predict the response to naturalistic images^{10,11}. Nevertheless, working with natural images directly is not always possible due to their complexity regarding critical parameters of signals processing, including visual content and its variability. The latter makes it hard to control the motion information that is effectively being presented to the visual system^{12,13}.

In the visual cortex, motion is detected at different levels and by populations of cells attuned to different parameters of the stimulus (see⁴ for a review). It has been reported that visual neurons stimulated with complex images, which should be closer to natural images, change their tuning curves compared to the response obtained with simple stimuli, becoming narrower and thus more selective to speed¹⁴. In turn, stimulation with naturalistic images generates sparser code^{15,16} and eye tracking responses become more precise¹⁷.

In the retina, the most basic form of motion detection is the response to moving textures in which Y-like *Retinal Ganglion Cells* (RGCs) increase their firing rate when sub-regions of their *Receptive Field* (RF) alternate between light and dark luminosity levels through time, regardless of the direction of motion^{18,19}. It has been demonstrated that these mechanisms arise from the RF properties of ON and OFF *Bipolar Cells* (BCs) that connect to the RGCs²⁰ and transmit their activity in a non-linear fashion. This non-linear integration of the activity of BCs allows that the alternating activation of cells of both polarities do not cancel each other upon reaching RGCs²¹. These BCs can generate fast transient responses when there is a luminance change of the corresponding polarity (dark to light or light to dark respectively), so each time that BCs detect a temporal change in luminance,

¹Centro Interdisciplinario de Neurociencia de Valparaíso, Universidad de Valparaíso, Valparaíso, Chile. ²Institut de Neurosciences de la Timone, Aix Marseille Université/CNRS, Marseille, France. ³Universidad Técnica Federico Santa María, Departamento de Electrónica, Valparaíso, Chile. Correspondence and requests for materials should be addressed to C.R.R. (email: cesar.ravello@cinv.cl) or M.-J.E. (email: mariajose.escobar@usm.cl) or A.G.P. (email: adrian.palacios@uv.cl)

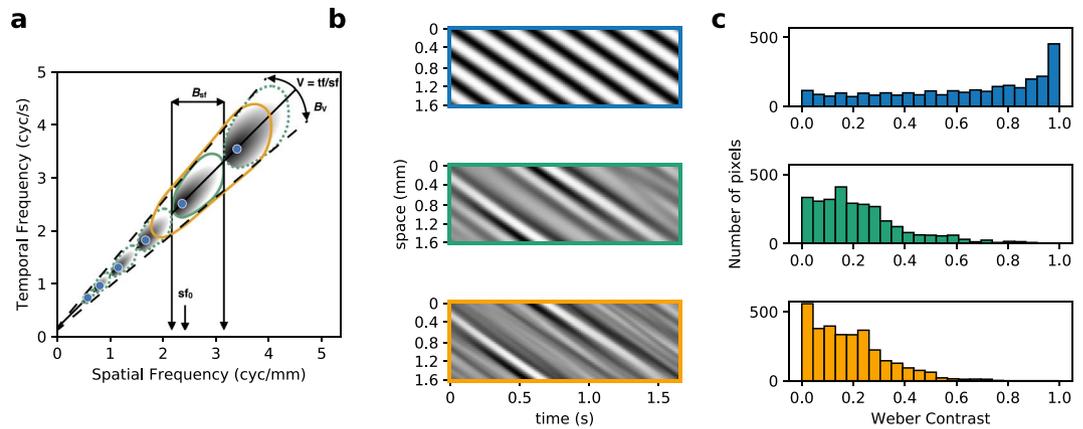


Figure 1. Motion Cloud (MC) stimuli are characterized by the parameters of their spatio-temporal spectral envelope. **(a)** Bi-dimensional representation of the spatio-temporal frequency space. Points along the continuous line correspond to simple drifting grating stimuli moving at a given speed but with different spatial frequencies (blue dots). MC stimuli have their spectral energy distributed in an “ellipse” warped along a given speed plane. Interestingly, these can be parameterized with the same mean spatial (sf_0) frequency and speed (V) as gratings, but with different levels of spatial frequency bandwidths (defined by the parameter B_{sf} with $B_{sf} = 0$ for the grating) which will define the level of complexity of the sequence. The green ellipse represents a narrow bandwidth stimulus and the orange ellipse a broader bandwidth stimulus. **(b)** Examples of stimuli at different complexity levels. For simplicity, we plot the light intensity along a single row of the image (vertical axis) for the whole duration of the stimulation (horizontal axis). We show the drifting grating (top panel), which is constituted by a single spatial and temporal frequency (and thus a single speed, seen as the slope in this view), and the MCs (middle and lower panels), which have the same speed and central spatial frequency but with a narrow or wide bandwidth in spatial frequency space. **(c)** Contrast distribution of the stimuli. The three types of stimuli were built to have the same Michelson contrast, but their contrast distribution is different. By construction, sinusoidal gratings have a higher proportion of brightest and darkest pixels, while the rest of the values have uniform representation; on the other hand, MCs have a large proportion of pixels with intermediate luminance, with extreme values being rare, resulting in images where low contrast is more frequent than high contrast, as is the case of natural images.

they will generate EPSPs in the corresponding RGCs. The integration time of the RGCs is longer than the time course of the response of the BCs, so fast successive activations will be added. In this way, the RGCs will respond with higher firing rates to higher speeds because of the higher rate of alternating activation of the ON and OFF sub-fields.

More specialized computations are carried out by, e. g. *Direction Selective RGCs* (DSRGC) which respond preferentially to a single direction in a varied range of speeds depending on the RGC type and the species^{4,3,22}. The circuitry involved in the computation of DSRGCs relates the integrated activity of BCs and starburst amacrine cells (SACs). At least two types of calculations are simultaneously present in the retina projecting to different places of the central nervous system^{23–25}. ON-OFF DSRGCs have been reported both in mouse and rabbits, responding to ON and OFF stimuli and being selective to four different types of motion direction in a broad range of speed. Besides, there also exists an On-DSRGC type reacting to slow velocities²³. Rabbit and salamander retina report *Object Motion* RGCs detectors, which fire when a small patch of the image moves in a different pattern than the rest of the background^{26,27}. More recently, it has been reported that the mouse retina contains a non-DS high-definition RGC, which is mainly selective to the motion of small objects across their RF²⁸. Another sophisticated processing of motion, present in salamander retina, is the response to *Motion Onset*²⁹, in which the initiation of movement within the RF of the RGC causes a higher response than does an object moving smoothly through it. While it has been shown that the retina is capable of advanced computations beyond standard processing³⁰, many capabilities remain unexplored³¹, for example, the fine tuning of retinal responses to motion in natural images.

In this study, using retinas from *Octodon degus*, a diurnal crepuscular rodent with a high proportion of cones (30%)³² and a relatively high density of RGCs, we focus on the properties of RGCs responses to motion in several scenarios. To overcome the limitations of working with simple, low-dimensional artificial stimuli, while at the same time avoiding the problems of working with natural images¹², we used synthetic random textures called Motion Clouds (MC)³³, which preserve some of the properties of natural images. The MCs mimic the spatiotemporal changes in luminance of natural scenes using a generative model for the geometric transformations of visual objects (translation, rotation, zoom)³⁴. In contrast to a traditional drifting grating which has a single spatial and temporal frequency, a MC resembles a dynamic texture for which components (textons) may have different sizes and move at different speeds (see Fig. 1a for comparison of these two stimuli). While in general, natural images have their power distributed along a broad range of frequencies, MC can be parametrized to have a restricted range of frequencies to limit their information content or to have a broad spectrum to resemble natural images. As such, each generated texture corresponds to a compound of different moving gratings with random positions (phases) yet with spatiotemporal properties distributed along a parameterizable area in Fourier space. Here, we

use the same distribution near the spatiotemporal speed plane, and control the amount of components by the bandwidth around the central spatial frequency. This defines a simple parameter to control the complexity of the input stimulus as the amount of compounded components in the texture (see Fig. 1b). This manipulation has previously been used successfully to study speed discrimination in humans¹⁷, where it was shown that additional information contained in the stimuli improves eye pursuit gain and precision.

To study such a property at the level of the retina, the activity recorded from a population of RGCs was compared using simple moving stimuli (drifting gratings) vs. MCs with the same speed and two levels of naturalness, regarding their spatial frequency spectrum (narrow and broad bandwidth). We found that for a significant fraction (66%) of RGCs responding to the motion, the cells' tuning bandwidths become narrower when the spatiotemporal frequency content of the stimulus increased. The set of RGCs studied contains cells traditionally assigned to different types, indicating that this property is not exclusive of a single class, but instead is widespread in the retina. At the population level, this change in tuning bandwidth results in sparser codes, which in turn reflects a more efficient coding of the motion information. Therefore, our findings show that fine-tuning of motion detection to natural image statistics already emerges at the level of the retina.

Results

We characterized the response from a population of *speed responsive* RGCs, recorded from retinal patches of 2 young *Octodon degus*, a diurnal rodent, having a total of 308 RGCs (see Methods section and Fig. 2 for details).

Speed responsive cells in the retina as characterized using gratings. We measured RGCs responses to a set of drifting gratings (artificial simple stimuli), with varying speed and spatial frequency (see Fig. 3a). For each spatial frequency, the response to speed averaged over trials was fitted to a skewed Gaussian. In Fig. 3b, we show how preference for certain speeds changes as a function of spatial frequency. As expected for RGCs, preferred speed decreases at higher spatial frequencies. In the example shown, the response is maximal at intermediate speeds and spatial frequencies, diminishing gradually as one moves away from that preferred combination of parameters. When looking at the fitted curves, it can be seen that the whole curve shifts towards lower speeds when increasing the spatial frequency.

A subset of the complete set of valid RGCs was separated and classified as *speed responsive cells* (SR), *i.e.* cells whose response is modulated by speed, by keeping only cells whose average firing rate correlated with the response strength, by estimating its correlation with the standardized component of the amplitude spectrum zFI ³⁵; and then, cells whose speed response curve to the drifting gratings has a good fit to eq. 3 at every spatial frequency tested ($\chi^2 < 0.05$ for the normalized response). We analyzed retinal patches from two animals, where 81 and 109 SR cells were found, respectively, representing $\approx 26\%$ and $\approx 29\%$ of the total RGCs recorded. Figure 4a shows the total number of RGCs found in a sample piece of retina, and their respective RF, of which the SR cells are highlighted in color (a second example is shown in Figure S1 in the Supplementary Information section).

To discriminate between SR and Non-SR cells, we analyzed the RF properties of each class. In terms of ON/OFF RGCs distribution, and similar to previous studies performed in the same animal model³⁶, a low ratio of ON/OFF cells were found. The temporal profile of a cell's response was characterized using the parameters indicated in Fig. 4b: zero-cross and peak-time. In Fig. 4c we show the distribution of these two parameters along the entire population of RGCs, separated by their speed responsive selectivity (SR cells are represented in red, while Non-SR in black). Interestingly, only the zero-cross parameter is significantly different between SR and Non-SR cell populations ($p = 0.026$ Kolmogorov-Smirnov test, $p = 0.65$ for the difference in peak time). Additionally, there was no direct match with other traditional criteria for classification of RGCs, such as Fast-Slow and Transient-Sustained using the latency and transience index³⁷ of the response to a full-field flash, since our set of SR cells contained cells belonging to each class (24 Fast-Transient, 13 Fast-Sustained, 19 Slow-Transient and 25 Slow-Sustained from 81 SR cells).

Speed selectivity changes with the complexity of the stimulus. When presenting a stimulus with the similar basic properties as the drifting gratings but with energy distributed along a larger area in parameter space, the responses from RGCs differ significantly. Examining in detail the response at the preferred spatial frequency (Fig. 5a) reveals that the firing rate response to lower and higher speeds is much weaker for the MC stimuli compared to the grating, while the responses become equivalent only around the preferred, intermediate speeds. This effect is more evident when looking at the tuning curves (Fig. 5b), where the response decreases more sharply when moving away from the preferred parameters for the complex stimuli. The change in the shape of the curve is evidenced by a change in the σ value of the fit, which decreases progressively from 1.76 to 1.45 and 0.95 as the complexity of the stimulus increases, while maintaining the same preferred speed (1.47; 1.37; 1.49 mm/s respectively). The firing rates are stable across repetitions, so the narrower tuning is not due to drifts or decay in cell activity. When considering all the conditions tested, the response profile to the MC stimuli cover a smaller area in parameter space along both axes, when viewing it as a two dimensional map of the response at each combination of parameters (Fig. 5c). As expected, speed preference decreases with spatial frequency, as seen by the slanted response profile.

This change in the response results in tuning curves with narrower profiles, thus the tuning of the RGC becomes more specific for certain parameters of the stimulus. Besides this change in the shape of the tuning curve, for many RGCs, the response to the MC stimuli has lower firing rates, as is the case for the example shown in Fig. 5b, while for others, the responses are equivalent and even higher in some cases (see examples in Fig. 5d).

Another interesting and unexpected finding is that some of the RGCs with narrower tuning for speed also present changes in their preferred speed, mostly towards higher speeds (Fig. 5d, panels on the right). Low firing rates obtained for MCs compared to gratings could be explained by differences in the contrast stimuli, but not the sharpening of the speed-selective curve (see Box 1). The low contrast observed in MCs is due to an increase of the

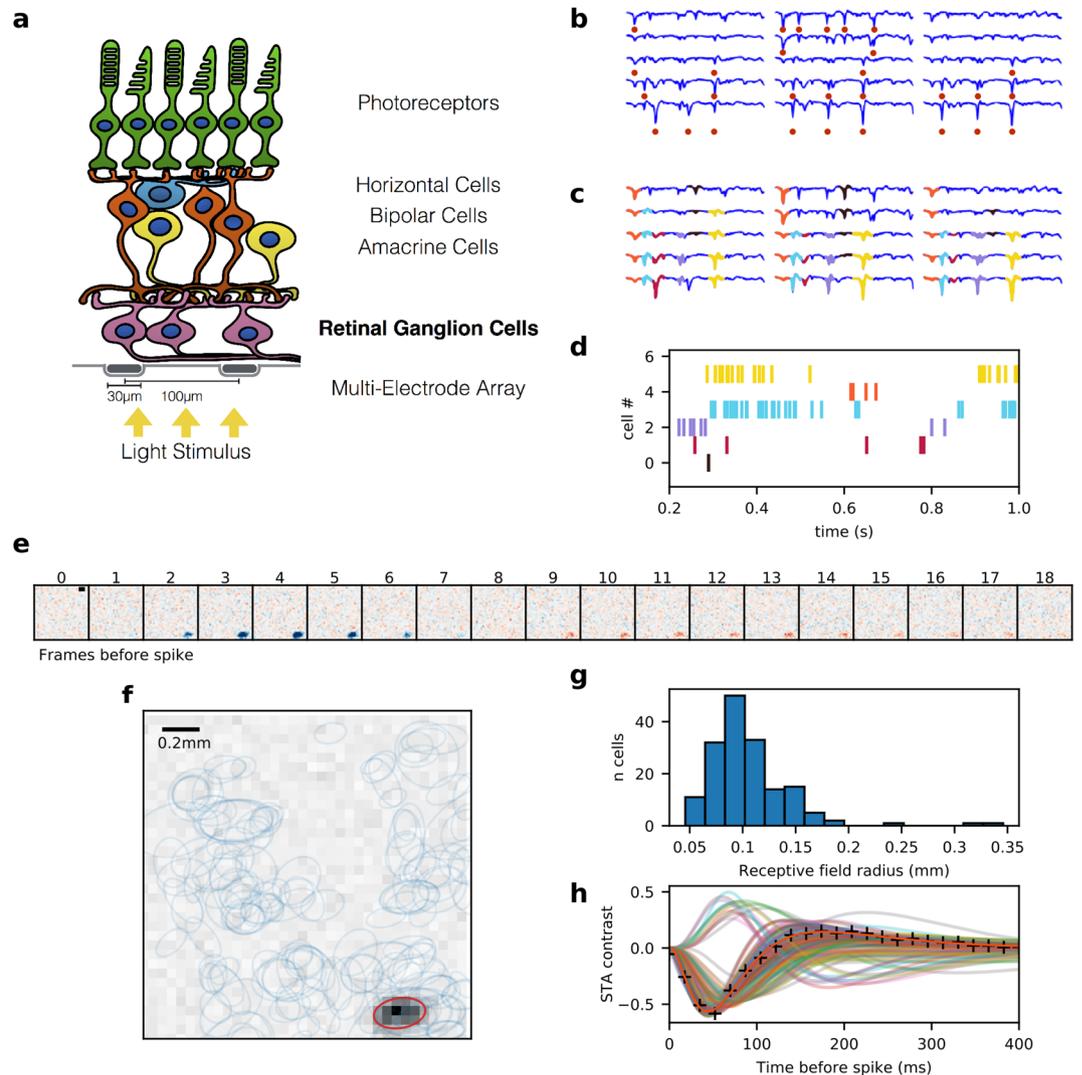


Figure 2. Experimental multielectrode recording setup. **(a)** Schematic view of the retina over a Multielectrode Array, while light stimulation is projected from below. **(b)** Each blue line represents a voltage signal from an electrode, and the red dots indicate detected peaks. **(c)** The signal of each electrode is reconstructed by iteratively adding different templates. Each color represents a different template, indicating spikes coming from different cells. Each spike can be detected by many neighboring electrodes, but only one will be assigned to a single unit. **(d)** The raster plot indicates the times at which each RGC fired an action potential. **(e)** Spike Triggered Average (STA) of a representative cell, computed from the response to checkerboard stimulus. Each panel shows the average image at the corresponding frame before the spiking of the cell (time zero). Blue pixels represent points with negative contrast, while red represent those with positive contrast. The black bar is 0.2 mm long. **(f)** Spatial component of the STAs of all the RGCs recorded from a retinal patch, after discarding non-valid units (see Methods for details). Each ellipse corresponds to a 2-D Gaussian fit at 1 s.d. The background shows the frame from (e) of maximum response and the red ellipse its fit, while the blue ellipses show the fit for the rest of the cells. **(g)** Distribution of RF sizes from a retinal patch, measured as the radius of a circle with the same area as the ellipse fit. **(h)** Temporal components of the STA of the cells shown in (g). The crosses mark the STA of the cell shown in (e).

spatial frequency bandwidth. In the grating analysis, we observe that for higher spatial frequency bandwidths, the peak activity for the central frequency gets smaller, and not narrower.

To quantify the change in the tuning bandwidth of the grating versus MC response, we measured the properties of the curve fitted to the response to speed (see Methods section for details). In Fig. 5d we show some examples of RGCs with narrower tuning in terms of lower response to the non-preferred speeds. At the population level, a large proportion of the SR cells show a decrease in the bandwidth of the response to speed at their preferred spatial frequency, for the complex stimuli when compared to the response to drifting gratings. This is evident when looking at the comparison of σ_v in Fig. 5e; for both MC series, a large portion of the points falls below the line of identity, meaning that the values are lower than for the grating stimulus. The distributions of σ_v are skewed towards lower values (Fig. 5f), meaning that a large portion of the speed responsive RGCs have

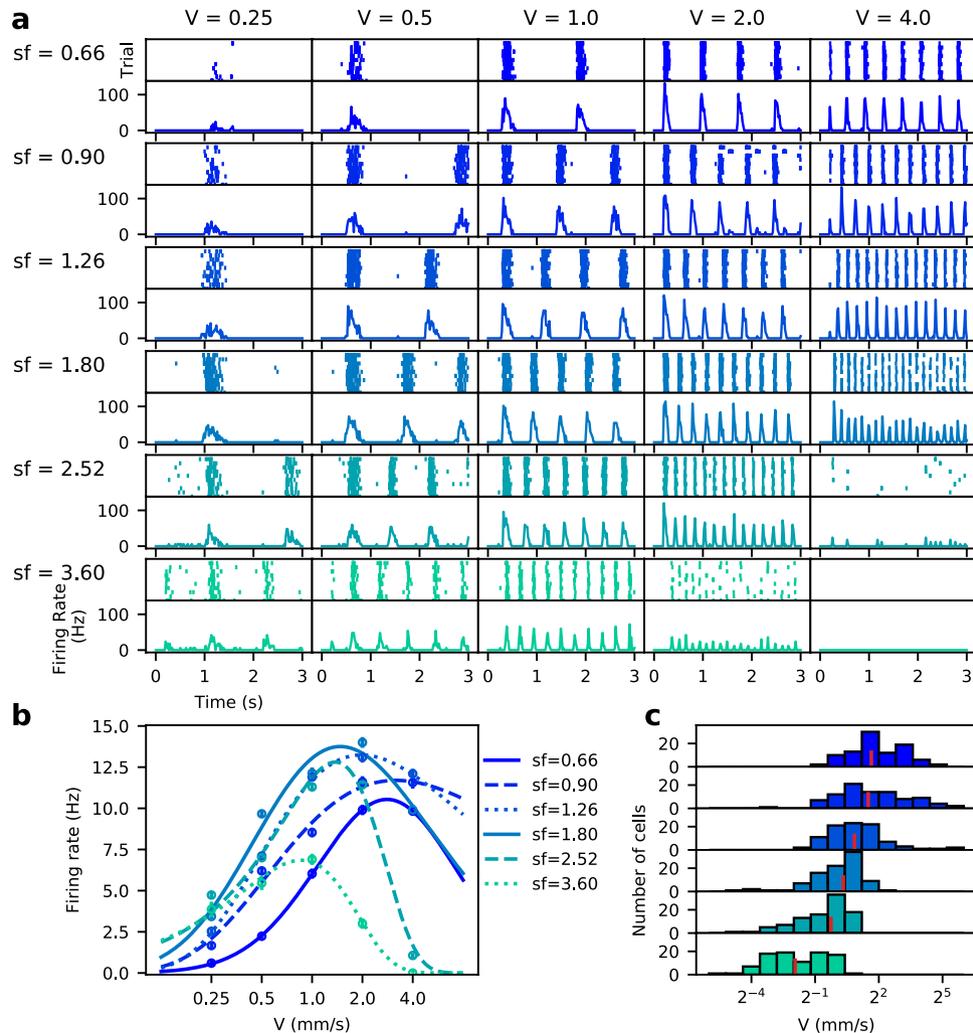


Figure 3. Response to variations in speed and spatial frequency of the drifting grating stimulus. **(a)** Raster and PSTHs of the response of a representative RGC at each speed and spatial frequency tested. **(b)** The response to speed at each spatial frequency sf was fitted to a skewed Gaussian. Each point is the average of the PSTH, with the error bars showing SEM. **(c)** All cells that had a good fit ($\chi^2 < 0.05$) at every sf were classified as *speed responsive cells*. For all these cells, the preferred speed at each spatial frequency is determined from the fits to their distribution. Vertical red lines show the median of the distribution for each condition.

narrower tunings when stimulated with the complex stimulus ($T = 1009$ $p = 0.00216$ for grating vs. narrow bw. MC, $T = 843$ $p = 0.00012$ for the broad bw. MC, Wilcoxon signed-rank test). With regard to the preferred speed, even though that for some RGCs the preferred speed increases (see the clusters at the lowest and highest speeds in Fig. 5g), overall, the differences are not highly significant (Fig. 5h $T = 1159$ $p = 0.018$ for grating vs. narrow bw. MC, $T = 1632$ $p = 0.893$ for the broad bw. MC, Wilcoxon signed-rank test).

Population responses becomes sparser for more naturalistic stimuli. When pooling the response of many RGCs, as a motion sensing neuron in higher cortical areas would do¹⁴, the response profile of the recorded population also shows differences when stimulated with the complex, more naturalistic stimuli. As expected from the individual responses, the response profile to speed depends on the spatial frequency (Fig. 6a left panel). The curve not only shifts when changing the spatial frequency, but the peak response also decays as the spatial frequency increases.

When comparing the response to the simple vs complex stimuli, the preference for lower speeds at higher spatial frequencies is still observed, but with a significant decrease in response for almost every condition (Fig. 6a middle and right panels). However, this decrease in response is larger for the lower speeds at every spatial frequency tested, to the point that for the two lowest speeds, the difference in response between spatial frequencies becomes non-significant (no significant difference at speeds 0.25 and 0.5 mm/s, $p > 0.05$ Wilcoxon signed-rank test), so the left part of all the tuning curves converge to the same values.

However, these lower firing rates elicited by the complex stimuli would not necessarily be detrimental to the coding properties of the retina. Instead, lower average firing rates could be associated with sparse coding. To see if this were the case, we computed the average population sparseness (eq. 5) for each stimulus condition (how many

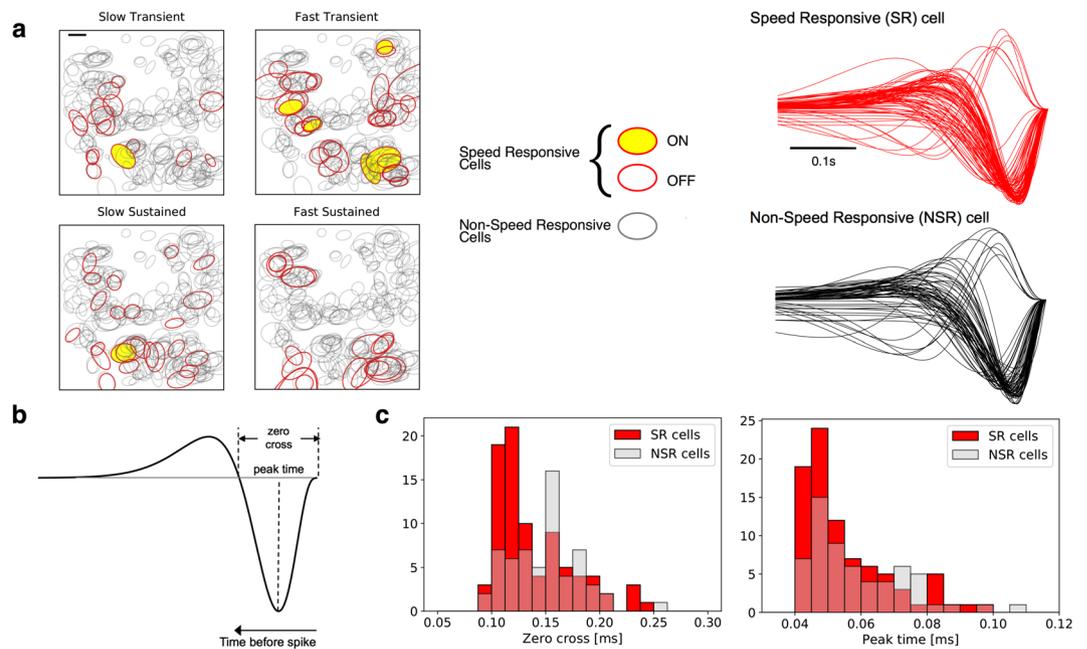


Figure 4. Speed responsive retinal ganglion cells. **(a) Left:** Map of the RGCs RF present in the experiment analyzed. RF of speed responsive (SR) RGCs (see text) are represented in color. **Right:** Temporal profiles of the receptive field center of SR (in red) and Non-SR (in black) cells. **(b)** Sample of a temporal profile showing the parameters extracted to compare SR versus Non-SR cells, which in this case are zero-cross and peak-time. **(c)** Histograms representing the distribution of the zero-cross (left) and peak-time (right) parameters in the SR and Non-SR population cells. We only observe significant differences for the zero-cross parameter ($D = 0.22$ $p = 0.026$ versus $D = 0.11$ $p = 0.65$ for the peak-time parameter, Kolmogorov-Smirnov test).

RGCs and how much they are responding to each stimulus). When compared to the simple stimulus, there are significant increases in sparseness in 8 conditions for the MC with narrow bandwidth and 19 conditions for the broad bandwidth stimulus at each combination of speed and spatial frequency tested (Fig. 6b).

A sparse response encodes motion information efficiently. To determine if the sparser retinal response is still encoding the motion information, we applied a reconstruction framework to decode the trajectory of a moving stimulus from the response of a group of RGCs and then extract information derived from the trajectory. As can be seen in Fig. 7, we constructed retinal images from the spikes emitted by each RGCs and its respective RF. To assess the ability to extract motion information from the reconstruction, we applied a process analogous to the way velocity is computed in cortical visual areas³⁸ (see Methods section for details). An example of this process is shown in Fig. 7. Even though that for some of the conditions the first stage of decoding was enough to get the correct speed (in the example, the lowest and the two highest speeds), in general the decoding improves at the Motion Energy and Opponent Motion stages, in the sense that the set of filters that produces the highest activation is the one that matches the presented speed.

The precision of speed decoding for the different conditions can be seen in Fig. 8. The motion traces obtained for all the stimuli can be seen in Fig. 8a. To measure the quality of the estimation, we computed the error rate as 1 minus the number of trials for which the estimation is correct, over the total number of trials. As seen in Fig. 8b, some of the conditions have zero error rate, while for the lowest speed, the error rate is 1 for many of the spatial frequencies for the three types of stimulus. For the gratings, most values are either one or zero, while for the MCs some conditions show intermediate values, meaning that for some trials the decoding was correct and for others it failed. When looking at the cumulative error rate across spatial frequencies (Fig. 8c) it becomes apparent that the decoding error for gratings is very high at the lowest and highest spatial frequencies, decreasing gradually towards the intermediate spatial frequencies. The same pattern can be seen for the MCs, however, the difference in decoding error for different spatial frequencies is lower, and stays within an acceptable range for all conditions. For broad bandwidth stimuli, the error rate is distributed more evenly across spatial frequencies, with bad decoding performance concentrated at the lowest speed.

While the distribution of the errors is different for the three types of stimuli, overall, the total error rate is slightly lower for the MC stimuli (Fig. 8c, rightmost plot, $T = 0.0$, $p = 0.016$ and $T = 3.0$, $p = 0.034$ Wilcoxon signed-rank test). Thus, it can be considered that there is a slight improvement in decoding performance. Since we were able to estimate the stimulus speed with an accuracy of ≈ 0.7 (1 - the total error rate) for the three types of stimuli, it could be argued from an information point of view, that the retina is transmitting at least enough information about the stimulus to encode speed under the three conditions. However, as shown in Fig. 5, MCs elicit responses with lower firing rates. So, if we quantify the cost of accuracy for each type of stimulus, in terms of the number of spikes used to successfully transmit the speed of the stimulus, we get that for the gratings the

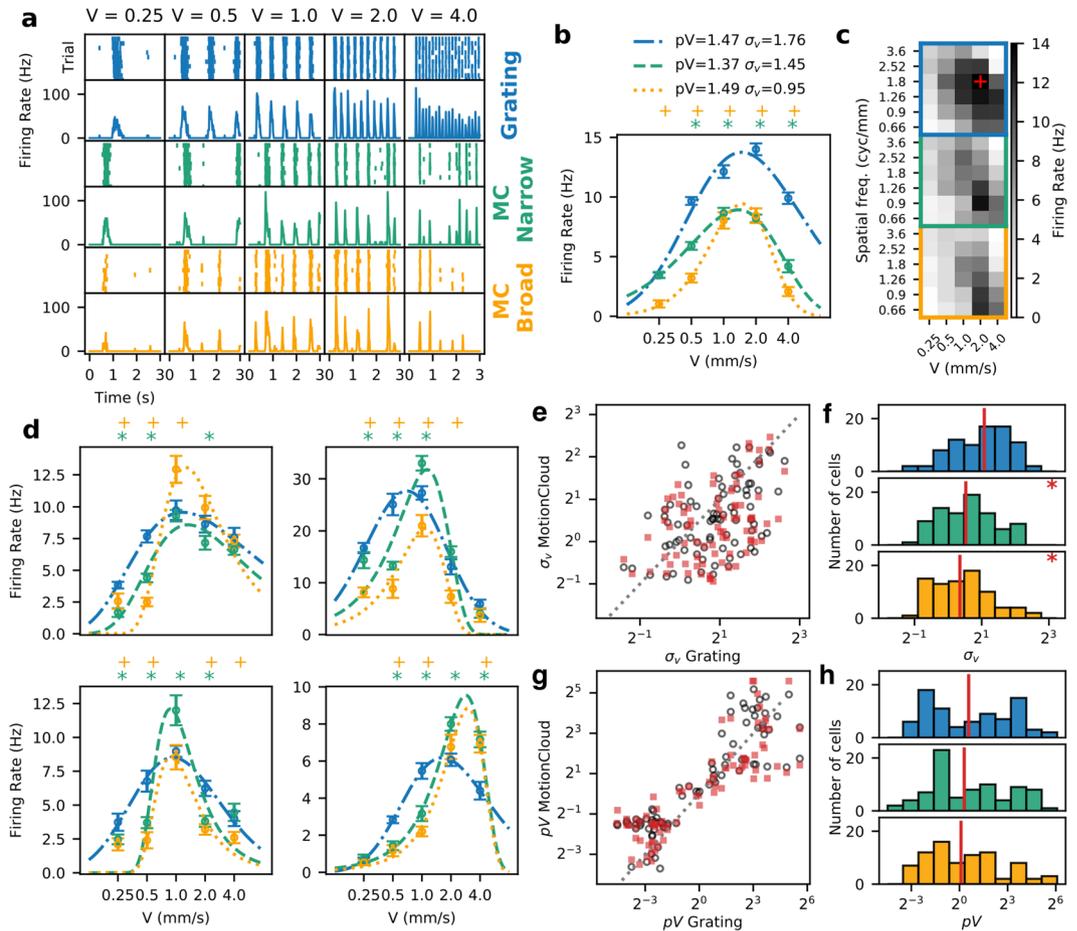


Figure 5. A large part of the population of observed RGCs narrow down their tuning selectivity with higher complexity stimuli. **(a)** Raster and PSTHs of the response of a representative cell for each type of stimulus and to the different speeds (in mms^{-1}) at its preferred sf_0 . **(b)** The change in tuning is evident when plotting the average response and its fit. Each point is the average of the PSTH, with the error bars showing s.d. The asterisks (for the Narrow bandwidth) and crosses (Broad bandwidth) mark the speeds at which the response to the MC is significantly different from the response to the grating ($p < 0.05$ Wilcoxon signed-rank test). The narrowing of the curve can be quantified by the decrease in the σ_v parameter, which in the case of this cell decreased with the complexity of the stimulus. **(c)** Spatiotemporal tuning of the cell for the different types of stimuli. The intensity at each square denotes the average firing rate for the corresponding stimulus parameters. The red cross indicates the point of maximum response, and corresponds to the spatial frequency plotted in panels a and b. **(d)** Additional examples of cells with narrower tuning for the naturalistic stimulus. **(e)** Joint distribution of speed tuning bandwidth coefficients σ_v of each Speed Responsive Cell. Empty circles show the relationship of grating against narrow bandwidth MC, filled squares show grating against broad bandwidth MC. In both conditions most of the points fall below the line of identity. **(f)** Histograms show the distribution of σ_v . Vertical red lines show the median of each distribution. The distribution for gratings is centered, while the distributions for both types of MC are skewed towards lower values with a significant difference ($T = 1009$ $p = 0.00216$ for grating vs. narrow bw. MC, $T = 843$ $p = 0.00012$ for the broad bw. MC, Wilcoxon signed-rank test). **(g,h)** same as **(e,f)** but for the preferred speed (peak of the fitted curve) with a less significant difference ($T = 1159$ $p = 0.018$ for grating vs. narrow bw. MC, $T = 1632$ $p = 0.893$ for the broad bw. MC, Wilcoxon signed-rank test).

value is ≈ 8.26 spikes/cell, while for the MC narrow and broad is 7.56 and 5.57 respectively, and these values are statistically significant ($T = 0.0$, $p = 0.005$ for the MC narrow and $T = 0.0$, $p = 0.0050$ for the MC broad, Wilcoxon signed-rank test).

Discussion

Our results strongly suggest that the tuning to naturalistic stimuli is already present at the level of the retina. For the first time, this showed that the tuning curves of response to motion become narrower under stimulation with increasingly more complex images, where we defined complexity as the number of compounded features of different spatial frequency but similar speed in the input texture. In particular, the bandwidth parameter controls complexity as it may generate a simple crystal-like stimulus (a grating component) with a very narrow bandwidth up to a smoke-like texture with a similar $1/f$ fall-off as natural images as the bandwidth grows to infinity. In the present work, the difference between each set of stimuli is the spatial frequency content, *i.e.* the variety in sizes for

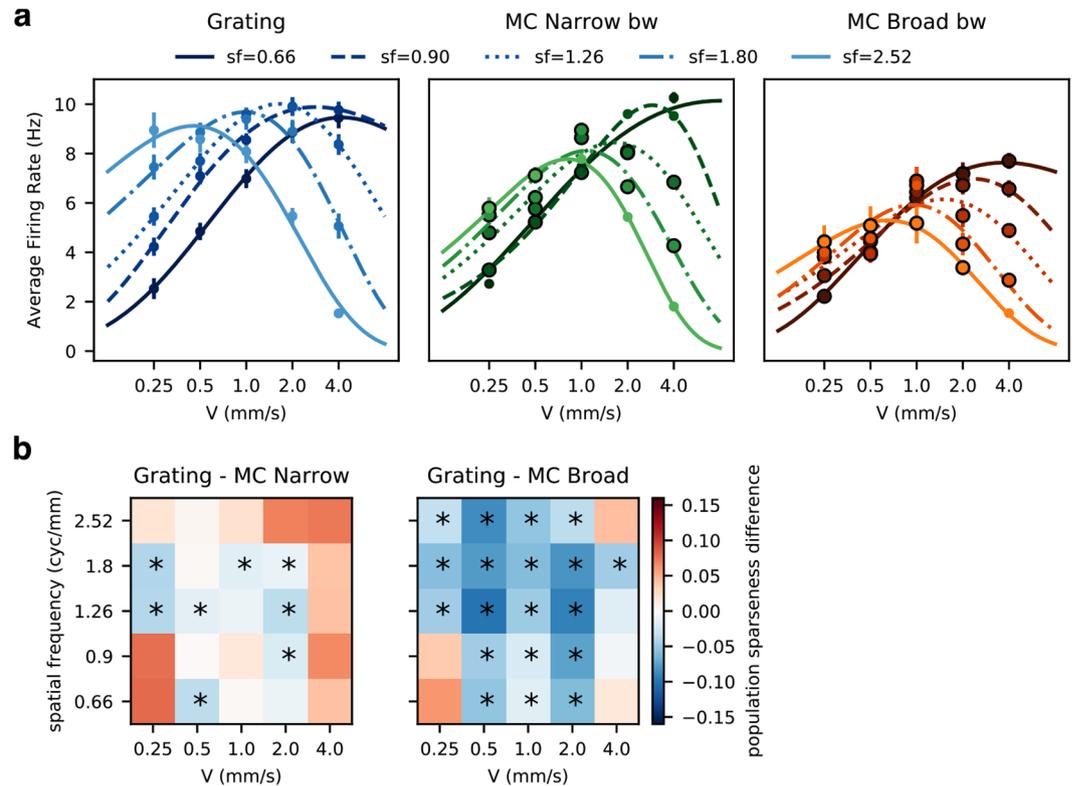


Figure 6. Naturalistic stimuli and sparse code. **(a)** Changes in Tuning to Speed at the population level. The dots show the average response across trials for the whole population of *Speed Responsive RGCs* for each combination of speed and spatial frequency used; the lighter shades show increasing spatial frequency, the error bars show standard deviation. The encircled dots indicate all the conditions for which the response is significantly smaller than the response to the grating stimulus ($p < 0.05$ Wilcoxon signed-rank test). The response to speeds for each spatial frequency was fitted to an skewed Gaussian. As expected, at higher spatial frequencies the preferred speed shifts towards lower values, together with a decrease in maximum response, however, this decrease becomes more pronounced when the complexity of the stimulus increases, to the point where the response to lower speeds becomes very similar for different spatial frequencies (no significant difference at speeds 0.25 and 0.5 mm/s ($p > 0.05$ Wilcoxon signed-rank test)). **(b)** Sparseness in the code, calculated as population sparseness (eq. 5), is a measure of how many RGCs are responding and their relative level of activity. The 2-D plots show the difference in average sparseness between grating and MC for each condition tested. In general, it is higher for the extreme parameters, however, for the more complex stimulus, the area of high sparseness is larger. The asterisks indicate the conditions for which the sparseness is higher compared to that obtained using grating stimulus (Wilcoxon signed-rank test, $p < 0.05$).

the elements that constitute the images; while the three classes of stimuli have the same average properties (spatial frequency and speed) for each set, the MC contain additional signals or components around this central value.

Additional signals in a stimulus can modulate the response of a RGC in different ways. It has been shown that isolated stimuli that fail to modify the activity of a RGC can alter the response to another stimulus when presented simultaneously³⁹. The latter would stem from nonlinearities in the processing of the signal performed by the RGC, in the sense that the response to a complex stimulus is not merely the response to the sum of its components. In some cases, the presence of signals around the RF, product of an increased area of stimulation, is enough to trigger changes in the response¹⁵. One of the mechanisms responsible for this effect is surround suppression, in which the presence of a stimulus in an area surrounding the spatial RF will decrease the response to the stimulus in the center, which is partly explained by an increase in the amplitude of inhibitory post-synaptic potentials¹⁶.

In our experiments, all stimuli cover the same area, which exceeds the area of the RFs of the recorded RGCs, so the presence of a stimulus in the surround would not by itself be enough to explain the changes in the observed response. Surround suppression occurs under a variety of stimulation patterns, both spatially and temporally⁴⁰. However, for simple stimuli it has been shown that the effect is most active when the parameters of the stimulus on the center and the surround are the same⁴¹, which would be the case for the gratings in our experiments; however, our data shows that this is not the case Fig. 5.

Seminal works in vision psychophysics established that spatial frequency sensitivity is dependent on contrast and overall luminance levels⁴². In general, sensitivity to high spatial and temporal frequency decreases at lower contrast levels. It could be argued then that our results reflect the effects of changes in contrast and not changes in spatial frequency. With our naturalistic stimulus, the relative contrast of each component of the images indeed scales with its spatial frequency, so components with higher spatial frequencies have lower contrast, following

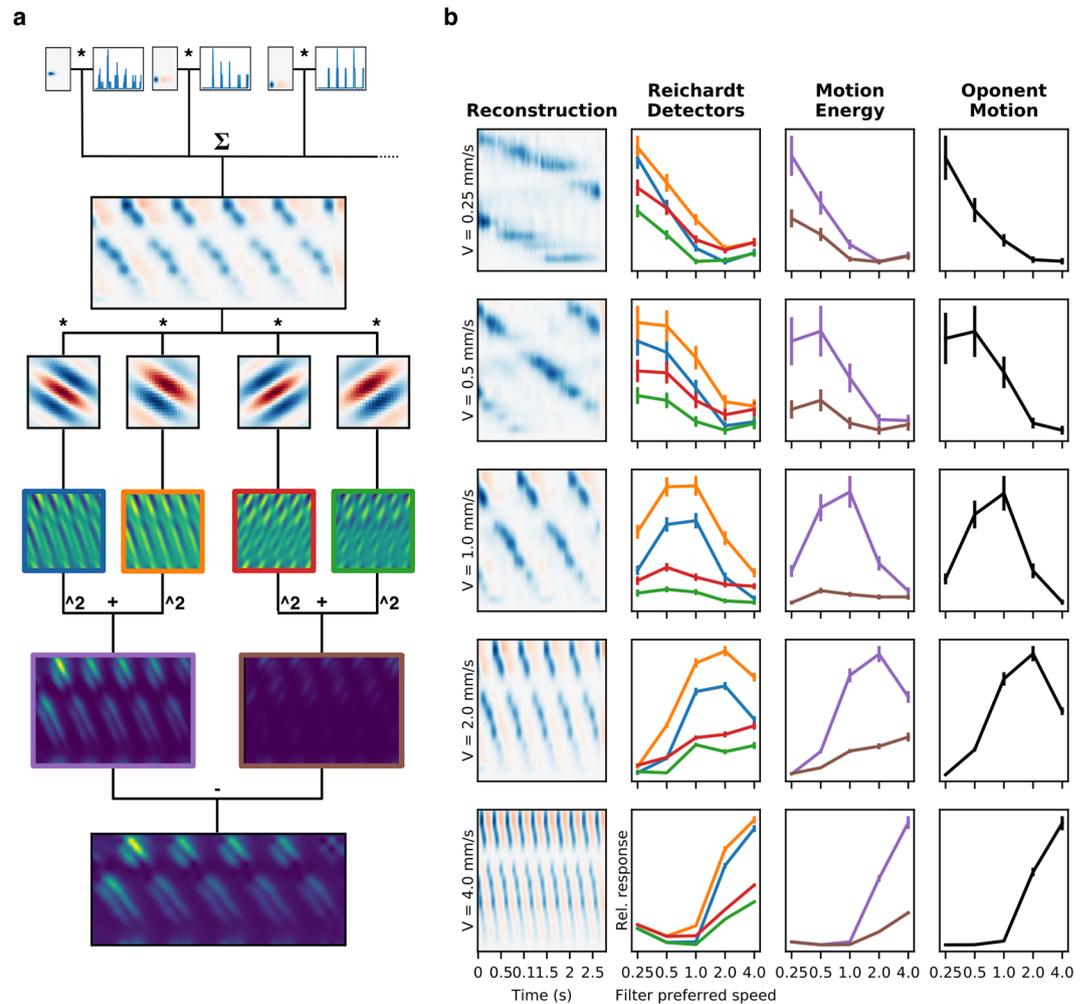


Figure 7. Reconstruction and decoding of motion stimuli. **(a)** The scheme shows the the stages of the motion decoding algorithm and the results of each step. **(b)** The leftmost column of each plot shows the reconstruction of the moving stimulus at different speeds, at a spatial frequency of 0.9cycles/mm. From the population response, the stimulus is reconstructed by convolution of the spike train with the RF of each RGC. The reconstructions are dominated by negative contrasts, due to the low number of ON type cells and their relatively low response. The mostly white row in each reconstruction corresponds to an area of the patch with no recorded units. The following columns show the response at each stage of decoding, for each corresponding reconstruction, averaged across trials. Vertical bars show s.d. The first stage is the response of each of the four filters that constitute the Reichardt detectors, the following column shows the Motion Energy in each direction and the rightmost column is the difference between the two. The color of each line corresponds to the respective panel of (a); For each reconstruction the decoding is considered correct if, in the final stage, the preferred speed of the most activated filter matches the stimulus speed.

the same $1/f$ scaling characteristic of natural images. However, the contrast of the main components (those with the central spatial frequency sf_0 or close to it) have the same contrast as the grating stimulus, while only the high-frequency components have lower contrast; this distribution of contrasts is stationary during the whole experiment, so the primary response drivers have the same contrast across the three types of stimuli. On the other hand, the effect of lower contrast on tuning curves is the opposite of what we report, as stimuli with lower contrast generate broader tuning curves⁴³, at least in the visual cortex. Finally, it has been shown that adaptation can alter the relationship between spatio-temporal tuning curves and contrast⁴⁴. Even though retinal adaptation - occurring at multiple stages - to luminance levels is relatively fast (in the order of seconds), the stimulation time necessary to generate spatial frequency adaptation is longer, so we explicitly designed the stimulation protocol to avoid adaptation by limiting the exposure time to 3 seconds and interspersing the images with 1 second of blank, average luminance stimulation. The stability in the response during stimulation and across trials for the three classes of moving stimulus (Fig. 5) shows that we can consider the effect of contrast, concerning the impact of spatial frequency bandwidth, to be negligible.

It has been shown that a non-linear interaction between the center and surround parts of the RF in RGCs modulates the response to spatial patterns in a context-dependent manner and that this effect is more significant during naturalistic stimulation⁴⁵. We think that this effect can explain, at least in part, the results that we observe;

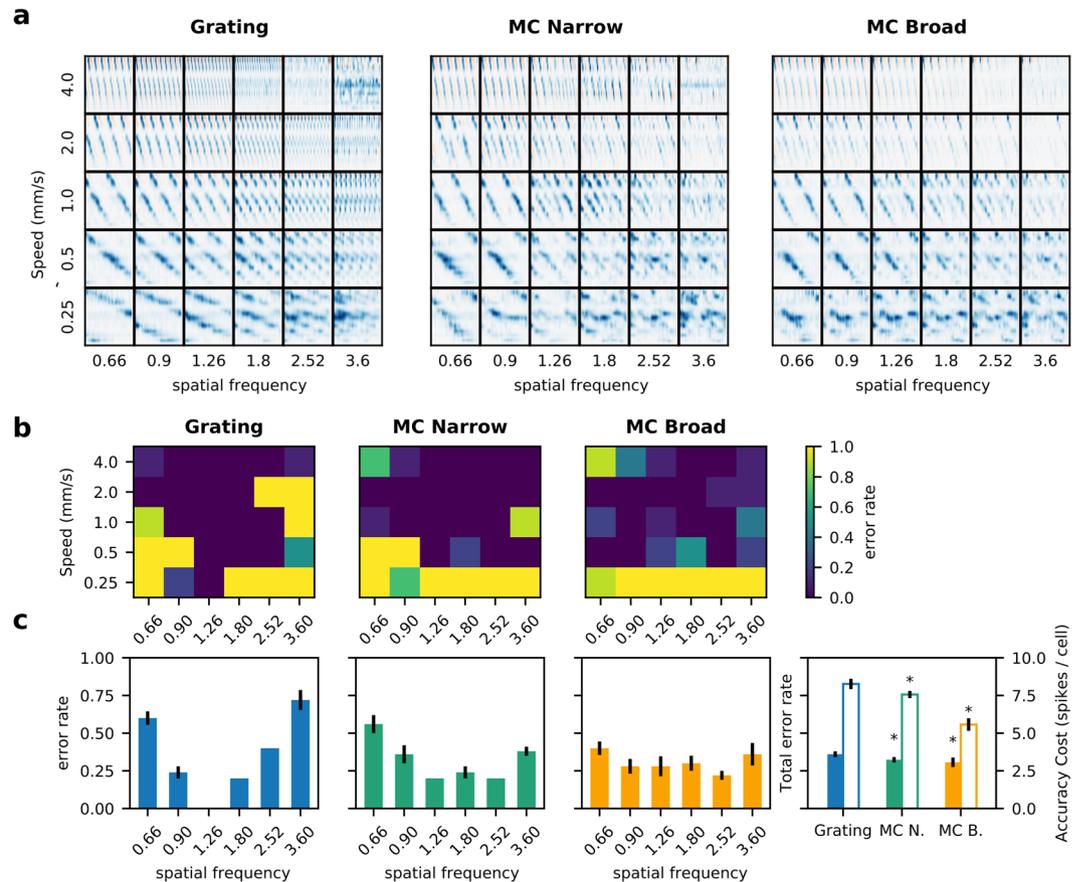


Figure 8. Precision of the estimation of stimulus speed. (a) Reconstruction of every motion stimulus tested. The representation is the same as in Fig. 7. For the three types of stimulus, the motion traces (and their slopes) are easily seen for all conditions, except at the lower speeds and higher spatial frequencies, which is consistent with lower responses under those conditions (Fig. 6). (b) Error rate in the decoding of the speed estimation (number of times that the algorithm estimated the correct speed over the total number of trials). The 2-D plots show the error rate at each spatial frequency and speed, for each type of stimulus. (c) Error rate for each spatial frequency and total error rate. Error lines show standard deviation between trials. Bars with no error mean that speed decoding was successfully performed between stimulus repetitions, for a given condition. In the rightmost plot, solid columns show total error rate and open bars show the Accuracy cost. Accuracy cost is computed as the success rate multiplied by the average firing (total number of spikes/number of cells); asterisks indicate significant difference with respect to the Grating ($p < 0.05$ Wilcoxon signed-rank test).

the additional signals contained in the complex stimulus would act as non-linear inhibitors, resulting in lower response when the parameters are farther from the preference of the RGC, producing narrower tuning curves. Moreover, the narrower speed selectivity tuning observed in RGCs, when stimulated with naturalistic-like stimuli such as MCs, seems to be a general property of many RGCs types in the retina and not related to a single RGC type. Recently, it has been reported the relevance of gap-junction connections between BCs in motion computation^{46,47}. This early mechanism, together with nonlinearity computations at the output of BCs⁴⁸, increases RGCs motion sensitivity to correlated spatio-temporal input, such as the case of MCs.

Interestingly, together with this narrower bandwidth, approximately a third of the RGCs show an increase in preferred speed (see two examples in Fig. 5, right panels), while others preserve their preference (Fig. 5, left panels). The change in the speed preference cannot be attributed to changes in the spatial frequency content of the stimulus. Due to the bandwidth of the stimuli, the images will include additional spatial frequencies, but the shape of the spectrum (see Methods: Naturalistic stimuli) determines that most of the additional frequencies will be higher than the central frequency (sf_0). If the RGC processed these higher spatial frequencies linearly, they should shift preference towards lower speeds - but our results show the opposite. From our analysis, it is difficult to relate or predict this behavior from other properties of the RGC's response, since the shift in the preferred speed appears not to correlate with any of the RF's characteristics.

Which are the possible implications or consequences of having these narrower tunings for naturalistic stimuli? At the level of individual RGCs, each RGC will only respond when the stimulus is near the preferred parameters. But this does not mean that a RGC will encode its preferred stimulus. Indeed, it has been shown that steeper tuning curves will encode more efficiently because stimuli will be more readily discriminated in the high slope range⁴⁹. As such, these results mean that for speed-selective RGCs an input with a broader spectrum of spatial

frequencies will have a higher precision in speed, *independently* of the local scale (spatial frequency) of the stimulus. This mechanism speaks to the dual mechanisms at the origin of selective responses (here speed) and of the invariance to other features (here scale). In addition to this effect at the individual RGC level, when looking at the population code, narrower tuning curves can lead to less overlap between RCGS, so in conjunction, the code will contain less redundancy, an essential aspect of the efficient coding theory^{50–52}. In support of this, we showed that population sparseness increases for the naturalistic stimuli, so in this respect, the retina is performing better under these types of stimulation, in the sense that a sparse code is more efficient from an information transmission point of view.

To see if these changes in tuning curves affect the coding of the motion signal, we implemented the stimulus reconstruction and speed decoding method. We chose speed as the relevant parameter for motion estimation, but we expect that with the proper adjustment, any property of the stimulus could be recovered, provided one can obtain a sufficiently good reconstruction⁵³. Another simplification that we took advantage of, is that our experimental design is based on the independent presentation of a single speed for each stimulus set (with slight variations around an average value for the MCs); this means that for the duration of each segment of stimulation, the whole population of RGCs is “seeing” the same speed. Thus, to recover the speed of the stimulus from the reconstruction, we only need to get a single value from the whole population and for the entire presentation of each sequence.

Our reconstruction method could be improved by varying some of the parameters, such as the weights of the sum, to minimize the error of the results, as is usually done in this type of work. However, even though when the models obtained by those methods are biologically plausible, the process of *a posteriori* maximization of the quality of decoding is hardly naturalistic, because no biological sensory system has access to the “real” value of the signal to perform the minimization step that is essential to this kind of approach.

Another possible improvement would be to incorporate nonlinearities in the decoding process, possibly by taking into account the history of the firing of a RGC instead taking every spike with the same value. Nonlinear models have been reported to give better stimulus reconstruction, especially for complex images⁵⁴.

In any case, even though our reconstructions have a relatively low degree of fidelity, they contain enough information to compute the speed of moving features in the images successfully. The latter is crucial because it means that even though these complex stimuli elicit lower retinal responses, the relevant information is still contained in these lower responses. In this regime of lower firing rates, energy expenditure would be lower, and the cost of information would be lower⁵⁵. Thus, this simplified approach is good enough to support the hypothesis that the retina would efficiently encode naturalistic stimuli.

In conclusion, we have shown that the retina, of a diurnal rodent, would be adapted to at least some of the characteristics of natural images, specifically those related to motion and the spatio-temporal information content and correlations used in this study. This adaptability is expressed in the narrower shapes of the tuning curves responses, which would lead to less overlap between RGCs in the feature space and thus less redundant and sparser population code, with smaller firing rates, which translates into less energy expenditure and less channel saturation. One limit of the stimuli that we used here is that they are densely packing the image space. However, natural images may exhibit some skewness and positive kurtosis in the distribution of luminance, reflecting the fact that the visual images are in general generated using a sparse number of components. This may introduce a different level of hierarchy in our definition of complexity: sparser textures may introduce spatial structures compared to the stationary nature of the MC textures used in this study. An interesting perspective is, therefore, to extend such stimulations in the retina to different levels of sparseness and to challenge the hypothesis that the processing in the retina could be differentially adapted to these different levels.

Methods

Animals. Young *Octodon degus* born in captivity are maintained in a controlled facility. Prior to each experiment, animals were put in darkness for 30 minutes, then deeply anesthetized with halothane and beheaded. Eyes were removed and dissected at room temperature under red illumination. Experimental procedures were approved by the Institutional Committee on Bioethics for Animal Research from the University and in accordance with the bioethics regulation of the Chilean Research Council (CONICYT).

Degus are diurnal rodents with 30% of their photoreceptors being cones and a comparatively high number of RGCs, qualities that makes them a good model for studying vision. In addition, the large area of the retina makes *degus* especially suitable for Multi Electrode array recordings as good coverage of all the electrodes in the matrix can be obtained. For a complete description of the model see^{32,56}.

Electrophysiological recordings. The experimental protocol was described before and follows³⁶. Briefly, the physiological response of *degu* RGCs to different types of visual stimuli was measured using a Multi Electrode Array^{57,58} (USB MEA256, Multichannel Systems GmbH, Reutlingen, Germany) using 256MEA100/30iR-ITO matrices and sampling at 20 kHz. After removing the eyes, the posterior hemisphere was dissected in quadrants and the pigmented epithelium separated from the retina. Finally, a piece of retina was mounted on a dialysis membrane and mounted on a perfusion chamber, which was then lowered onto the electrode array with the RGCs side down. Recording commenced under perfusion with AMES medium bubbled with 95% O₂ 5% CO₂ at 33 °C and the pH adjusted to 7.4.

Visual Stimulus Presentation. Stimulus display was performed with a conventional DLP projector using custom optics to reduce and focus the image onto the photoreceptor layer, projecting from the RGC side. The projected size of a pixel is $\approx 4 \mu\text{m}$, maintaining an average irradiance of 70 nWmm^{-2} , covering an area of $\approx 2 \times 2 \text{ mm}$, exceeding the area covered by the electrodes. Regarding the spectral emission of the DLP used here, the LEDs with peak emission at 460 nm and 520 nm (USB4000 Fiber Optic Spectrometer, Ocean Optics) covers the main

green-cones with peak at 510 nm and barely, to -2 log maximal peak sensitivity of, the 370 nm UV-cones present in the dichromatic visual system of *Octodon degus*⁵⁶. Cones in degus are close to a 30% of total number of photoreceptors where green-cones (or M-pigment) represent 85% and UV-cones 15% (or UV-pigment)³². In that respect our main results and conclusions, at photopic conditions, are extracted from the response of M-pigment cones. Stimuli were generated using the code available at <https://github.com/NeuralEnsemble/MotionClouds>. Timing of the images was controlled using a custom-built software based on Psychtoolbox for MATLAB. Spike sorting analysis was performed using *Spyking-Circus*⁵⁹. Data was analyzed via redistributable Jupyter notebooks using SciPy and running Python 3 kernels. All data fitting procedures were performed with LMFIT. Figures were generated with Matplotlib.

Maximal speed projection is given by the time needed for a point to cross half of the projected image in one frame time, i.e., 200 px/frame. As the projector works at 60 frames, it gives us an estimation of 12000 px/s. Considering that one pixel covers $4\ \mu\text{m}$ ($4\ \mu\text{m}/\text{px}$), it generates a maximal speed of 48 mm/s, which is higher than the highest rate used in our study (4 mm/s).

Characterization of spatiotemporal tuning of RGCs. To produce a standard characterization of RGCs responses to different stimulations, we measured the response to a white noise checkerboard pattern (block size = 0.05 mm, 35×35 blocks) for 1200 s at 60 Hz and to sinusoidal drifting gratings at maximum contrast (minimum and maximum Weber contrast of ≈ -94 and ≈ 129 respectively), with spatial frequencies of 0.66, 0.9, 1.26, 1.8, 2.52 and 3.6 cycles/mm, and speeds of 0.25, 0.5, 1.0, 2.0 and 4.0 mm/s in sequences of three seconds with 10 repetitions of each. Due to the constraints in recording length stemming from tissue viability, we focused only on changes in response to the speed of the moving stimulus and not to its direction, so the present protocol uses a single direction for all sets of stimuli.

RFs were estimated from the response to the checkerboard stimulus by reverse correlation, yielding the *Spike-Triggered Average (STA)*⁶⁰ as a three-dimensional spatiotemporal impulse response (Fig. 2e). The spatial characterization was performed by fitting a bi-dimensional Gaussian function at the time point of maximum amplitude, then drawing an ellipse at one standard deviation. The size of each RF was calculated as the radius of the circle with the same area as the ellipse fit⁶¹. The shape was quantified by the ellipse-eccentricity ε defined as $\varepsilon = \sqrt{1 - (b/a)^2}$ where a and b are respectively the radius of the major and minor axis; since noisy STAs generate fittings with high ellipse-eccentricities, cells with $\varepsilon > 0.9$ were discarded. The temporal profile was computed as the time course of the intensity at the point with the largest variance, and then fitted to a difference of two cascades of low-pass filters⁶². Basal activity was set to zero and amplitude normalized, such that it follows the form

$$R(t) = p_1 \left(\frac{1}{\tau_1} \right)^n e^{-n \frac{t}{\tau_1 - 1}} - p_2 \left(\frac{1}{\tau_2} \right)^n e^{-n \frac{t}{\tau_2 - 1}} \quad (1)$$

where t represents time in number of image frames before the spike, τ_1 and τ_2 describe the temporal decay of the response of each filter; scalars p_1 and p_2 are amplitude responses of each filter, and n is a free parameter.

RGCs sensitive to motion. Tuning to motion was evaluated from the response to the drifting gratings; the response to each set of speeds and spatial frequencies was measured by *Peristimulus Time Histogram (PSTH)* and Fourier analysis for 10 repetitions of each sequence. The beginning of the PSTH is set to 200 ms after stimulus start to discard the response to stimulus onset. To select the cells responsive to motion, first we selected RGCs with response modulated by the stimulus. As a measure of the strength of the modulation of the response by the stimulus, we computed the *standardized F1 (zF1)*³⁵, defined as:

$$zF1 = \frac{F1 - \text{mean}(FFT)}{SD(FFT)}, \quad (2)$$

where $F1$ is the amplitude component at the temporal frequency of the stimulus, $\text{mean}(FFT)$ is the mean amplitude of the spectrum over the range of frequencies from $1/T$ to the Nyquist frequency, and $SD(FFT)$ is the standard deviation of amplitudes in the frequency spectrum over the same range of frequencies. We discarded all the cells whose $zF1$ values were not correlated with the corresponding average firing rate (Pearson correlation, $p > 0.05$).

Tuning to speed v was fitted to a skewed Gaussian in which the response to speed $R(v)$ takes the form

$$R(v) = A \left[\exp \left(\frac{-(\log_2 v - \log_2 V)^2}{2 \times (\sigma_v + \zeta \times (\log_2 v - \log_2 V))^2} \right) - \exp \left(\frac{-1}{\zeta^2} \right) \right], \quad (3)$$

where V is the preferred speed, σ_v the curve width, ζ represents the skew of the curve and A a scaling factor¹⁴. Since σ_v is highly correlated with tuning bandwidth (width of the function at half-height)⁶³, we used σ_v for all bandwidth related functions. Changes in tuning bandwidth were measured as the normalized difference in σ_v , as

$$\Delta \sigma_v = \frac{\sigma_{v1} - \sigma_{v2}}{\sigma_{v1} + \sigma_{v2}}. \quad (4)$$

From eq. 3, we define *Speed Responsive (SR)* RGCs as those whose response to the drifting gratings has a good fit to the equation at every spatial frequency tested ($\chi^2 < 0.05$).

sf (cycles/mm)	tf (cycles/sec)	V (mm/sec)	B_{sf} Narrow	B_{sf} Broad
0.66	0.165	0.25	0.24	0.66
	0.33	0.5		
	0.66	1.0		
	1.32	2.0		
	2.64	4.0		
0.9	0.225	0.25	0.36	0.9
	0.45	0.5		
	0.9	1.0		
	1.8	2.0		
	3.6	4.0		
1.26	0.315	0.25	0.54	1.26
	0.63	0.5		
	1.26	1.0		
	2.52	2.0		
	5.04	4.0		
1.8	0.45	0.25	0.72	1.8
	0.9	0.5		
	1.8	1.0		
	3.6	2.0		
	7.2	4.0		
2.52	0.63	0.25	1.08	2.52
	1.26	0.5		
	2.52	1.0		
	5.04	2.0		
	10.08	4.0		
3.6	0.9	0.25	1.5	3.6
	1.8	0.5		
	3.6	1.0		
	7.2	2.0		
	14.4	4.0		

Table 1. Parameters used to build each set of motion stimuli. For the gratings, the temporal frequency tf is modulated to obtain the same speeds at every spatial frequency sf . Motion Clouds (MCs) do not have a tf parameter, they are controlled directly by the speed parameter V . Note that the MCs software works in pixel units, so a proper scaling has to be applied. In our case, the conversion factor between pixels and mm is 240:1.

The selection criteria for SR RGCs did not elicit direction or orientation selective responses. Moreover, computing the Direction Selectivity Index for all the SR RGCs, we observe practically no cell with a DSI larger than 0.5, the value that is usually considered the minimum to classify a cell as Direction Selective.

Population coding. Population response was evaluated as the average of the PSTH of all *speed responsive* RGCs. Sparseness of the response to each stimulus was measured as Population Sparseness (S_p), defined in⁶⁴ as

$$S_p = \left[1 - \frac{(\sum_{i=1}^N |A_i|/N)^2}{\sum_{i=1}^N |A_i|^2/N} \right] \times \left(1 - \frac{1}{N} \right)^{-1} \quad (5)$$

where A_i is the average response of cell i to the stimulus.

Naturalistic stimuli. We used Motion Clouds (MCs)³³, with an image size of 400×400 px (projected size $\approx 1.6 \times 1, 6$ mm) for 3 s, to measure the response of the retina to the motion of stimuli with different levels of naturalness (in terms of their spatio-temporal correlations). MCs are synthetic dynamical textures that mimic some key properties of natural images³⁴, while allowing for the precise control over the signals that are being presented, which is a highly desirable property for the study of sensory systems under naturalistic conditions¹². The traits of these textures are determined by three parameters for the center of the envelope (size, speed, and orientation), along with their respective bandwidths. These bandwidths control the area (in parameter space) in which the power of the stimulus will be spread.

Our protocols consisted of a set of different sequences with the same motion information (mean speed and spatial frequency), but in which we progressively varied the spatial frequency bandwidth (*i.e.* the range of sizes), as this was reported to influence visual motion detection¹⁷. These match the parameters of the drifting grating stimuli (equivalent to an MC with infinitely narrow bandwidth), and textures with a narrow and a wide

bandwidth of spatial frequencies (Fig. 1a blue, green and orange ellipses respectively). For narrow and broad bandwidth textures, amplitude scales with spatial frequency: for the narrow bandwidth stimuli, the scaling factor is chosen for the spectra to not overlap, while for the broad band stimuli the scaling factor is 1 resulting in $B_{sf} = sf_0$, so that each spectrum overlaps with the adjacent one along the line of equal speed, see Table 1 for the specific parameters. Figure 1b shows an example stimulus, comparing the spatiotemporal image of a drifting grating and MCs with narrow and broad bandwidth, illustrating the smooth transition from a “crystal-like” pattern (the grating) to progressively more complex and naturalistic stimuli using a single parameter (B_{sf}). Under this characterization, a collection of well-sampled natural images would correspond to MCs with infinite bandwidth, *i.e.* it would contain every spatial frequency, so drifting gratings and natural images would be on opposite ends of the complexity scale, with our MCs constituting a gradual transition between them. The responses for the three types of stimuli were constructed to obtain a similar Michelson contrast (that is, their intensity values vary from 0 to 1). However, the intensity distribution was different, as seen in Fig. 1c. Intensity was quantified by measuring the Weber contrasts distribution: for each intensity value the average intensity was subtracted and divided by the average intensity. Because of their design, sinusoidal grids have a higher proportion of brightest and darkest pixels, while the rest of the values have a uniform representation. On the other hand, the MCs have a large portion of pixels with intermediate luminance, with extreme values being rare, which resembles natural images.

Trajectory Reconstruction and velocity estimation. The first step is the reconstruction of the trajectory of a moving stimulus from the response of a population of RGCs. As recently demonstrated in⁵⁴, the *Decoding Fields* obtained by maximum likelihood are very similar to the RFs obtained by traditional methods of reverse correlation. Based on this, we made a spatiotemporal reconstruction of the stimulus by convolving the response vectors of each RGCs with its respective RF and then performing a weighted sum across the whole population. The response r_i^c is the number of spikes at each interval i with $\Delta t = 16.66$ ms. The RF is computed by reverse correlation from the response to a checkerboard stimulus. Since we are interested in stimuli moving only in one direction, we collapse the tridimensional spatiotemporal RF to a bidimensional representation $RF_{x,t}^c$ containing only time and one spatial dimension. Thus, the intensity of the reconstructed stimulus at each point in space and time, $I(x, i)$, is given by

$$I(x, i) = a + \sum_c \omega^c \sum_{j=0}^{N-1} r_{i-j}^c RF_{x,j}^c \quad (6)$$

where a is a constant offset, ω^c is the weighting factor for each cell, $RF_{x,j}^c$ is the value of the receptive field at point x at time $j\Delta t$ before the spike, and $N\Delta t$ is the length of the RF estimation.

Speed of the moving stimulus is assessed by a multi-stage process following standard motion processing models³⁸. The first step is the convolution of the reconstructed stimulus with sets of filters resembling Reichardt detectors, approximated by slanted Gabor filters oriented in space time with preference for different speeds (given by the slope in the space/time plane, thus the speed tuning is given by the angle of the Gabor kernel). Each kernel is paired with another of the same properties, but with its phase shifted, forming a quadrature pair. The response of each quadrature pair is rectified and added to obtain the Motion Energy for each speed. The process is repeated to compute the Motion Energy in the opposite direction, and finally, both are subtracted to obtain the net motion signal for each speed. The estimated speed is determined by the set of filters with the highest activation in a winner-takes-all approach. An example of the process is shown in Fig. 7.

References

- Barlow, H. B., Hill, R. M. & Levick, W. R. Retinal ganglion cells responding selectively to direction and speed of image motion in the rabbit. *The Journal of Physiology* **173**, 377–407 (1964).
- Olveczky, B. P., Baccus, S. A. & Meister, M. Retinal adaptation to object motion. *Neuron* **56**, 689–700 (2007).
- Vaney, D. I., Sivy, B. & Taylor, W. R. Direction selectivity in the retina: symmetry and asymmetry in structure and function. *Nature Reviews Neuroscience* **1–15** (2012).
- Burr, D. & Thompson, P. Motion psychophysics: 1985–2010. *Vision Research* **51**, 1431–1456 (2011).
- Simoncelli, E. P. & Olshausen, B. A. Natural image statistics and neural representation. *Annual Review of Neuroscience* **24**, 1193–1216 (2001).
- Geisler, W. S. Visual Perception and the Statistical Properties of Natural Scenes. *Annual Review of Psychology* **59**, 167–192 (2008).
- Dyakova, O., Lee, Y.-J., Longden, K. D., Kiselev, V. G. & Nordström, K. A higher order visual neuron tuned to the spatial amplitude spectra of natural scenes. *Nature communications* **6**, 8522 (2015).
- Yu, Y., Romero, R. & Lee, T. S. Preference of Sensory Neural Coding for 1/f Signals. *Physical Review Letters* **94**, 108103–4 (2005).
- Olshausen, B. & Field, D. Sparse coding of sensory inputs. *Current opinion in neurobiology* **14**, 481–487 (2004).
- David, S. V. Natural Stimulus Statistics Alter the Receptive Field Structure of V1 Neurons. *Journal of Neuroscience* **24**, 6991–7006 (2004).
- Carandini, M. *et al.* Do we know what the early visual system does? *The Journal of neuroscience: the official journal of the Society for Neuroscience* **25**, 10577–10597 (2005).
- Rust, N. C. & Movshon, J. A. In praise of artifice. *Nature neuroscience* **8**, 1647–1650 (2005).
- Geisler, W. S. & Burge, J. Optimal speed estimation in natural image movies predicts human performance. *Nature communications* **6**, 1–11 (2015).
- Priebe, N. J. Tuning for Spatiotemporal Frequency and Speed in Directionally Selective Neurons of Macaque Striate Cortex. *Journal of Neuroscience* **26**, 2941–2950 (2006).
- Vinje, W. E. & Gallant, J. L. Natural stimulation of the nonclassical receptive field increases information transmission efficiency in V1. *The Journal of neuroscience: the official journal of the Society for Neuroscience* **22**, 2904–2915 (2002).
- Haider, B. *et al.* Synaptic and network mechanisms of sparse and reliable visual cortical activity during nonclassical receptive field stimulation. *Neuron* **65**, 107–121 (2010).
- Simoncini, C., Perrinet, L. U., Montagnini, A., Mamassian, P. & Masson, G. S. More is not always better: adaptive gain control explains dissociation between perception and action. *Nature neuroscience* **15**, 1–10 (2012).

18. Enroth-Cugell, C., Robson, J. G., Schweitzer-Tong, D. E. & Watson, A. B. Spatio-temporal interactions in cat retinal ganglion cells showing linear spatial summation. *The Journal of Physiology* **341**, 279–307 (1983).
19. Shapley, R. & Lennie, P. Spatial frequency analysis in the visual system. *Annual Review of Neuroscience* **8**, 547–581 (1985).
20. Demb, J. B., Haarsma, L., Freed, M. A. & Sterling, P. Functional circuitry of the retinal ganglion cell's nonlinear receptive field. *The Journal of neuroscience: the official journal of the Society for Neuroscience* **19**, 9756–9767 (1999).
21. Demb, J. B., Zaghloul, K., Haarsma, L. & Sterling, P. Bipolar cells contribute to nonlinear spatial summation in the brisk-transient (Y) ganglion cell in mammalian retina. *The Journal of neuroscience: the official journal of the Society for Neuroscience* **21**, 7447–7454 (2001).
22. Demb, J. B. Cellular Mechanisms for Direction Selectivity in the Retina. *Neuron* **55**, 179–186 (2007).
23. Kanjhan, R. & Sivyer, B. Two types of ON direction-selective ganglion cells in rabbit retina. *Neuroscience Letters* **483**, 105–109 (2010).
24. Wei, W., Hamby, A. M., Zhou, K. & Feller, M. B. Development of asymmetric inhibition underlying direction selectivity in the retina. *Nature* **469**, 402–406 (2011).
25. Cruz-Martín, A. *et al.* A dedicated circuit links direction-selective retinal ganglion cells to the primary visual cortex. *Nature* **507**, 358–361 (2014).
26. Olveczky, B. P., Baccus, S. A. & Meister, M. Segregation of object and background motion in the retina. *Nature* **423**, 401–408 (2003).
27. Baccus, S. A., Olveczky, B. P., Manu, M. & Meister, M. A Retinal Circuit That Computes Object Motion. *Journal of Neuroscience* **28**, 6807–6817 (2008).
28. Jacoby, J. & Schwartz, G. W. Three Small-Receptive-Field Ganglion Cells in the Mouse Retina Are Distinctly Tuned to Size, Speed, and Object Motion. *Journal of Neuroscience* **37**, 610–625 (2017).
29. Chen, E. Y. *et al.* Alert Response to Motion Onset in the Retina. *Journal of Neuroscience* **33**, 120–132 (2013).
30. Gollisch, T. & Meister, M. Eye Smarter than Scientists Believed: Neural Computations in Circuits of the Retina. *Neuron* **65**, 150–164 (2010).
31. Masland, R. H. & Martin, P. R. The unsolved mystery of vision. *Current Biology* **17**, R577–R582 (2007).
32. Jacobs, G. H., Calderone, J. B., Fenwick, J. A., Krogh, K. & Williams, G. A. Visual adaptations in a diurnal rodent, *Octodon degus*. *Journal of comparative physiology. A, Neuroethology, sensory, neural, and behavioral physiology* **189**, 347–361 (2003).
33. Leon, P. S., Vanzetta, I., Masson, G. S. & Perrinet, L. U. Motion clouds: model-based stimulus synthesis of natural-like random textures for the study of motion perception. *Journal of Neurophysiology* **107**, 3217–3226 (2012).
34. Vacher, J., Meso, A. I., Perrinet, L. U. & Peyré, G. Biologically inspired dynamic textures for probing motion perception. In Cortes, C. *et al.* (eds) *Advances in Neural Information Processing Systems* **28**, 1909–1917, <http://arxiv.org/abs/1511.02705> (Curran Associates, Inc., 2015).
35. Wypych, M. *et al.* Standardized F1: a consistent measure of strength of modulation of visual responses to sine-wave drifting gratings. *Vision research* **72**, 14–33 (2012).
36. Palacios-Muñoz, A. *et al.* Role of connexin channels in the retinal light response of a diurnal rodent. *Frontiers in cellular neuroscience* **8**, 249 (2014).
37. Farrow, K. & Masland, R. H. Physiological clustering of visual channels in the mouse retina. *Journal of Neurophysiology* **105**, 1516–1530 (2011).
38. Bradley, D. C. & Goyal, M. S. Velocity computation in the primate visual system. *Nature Reviews Neuroscience* **9**, 686–695 (2008).
39. Vilankar, K. P. & Field, D. J. Selectivity, hyperselectivity, and the tuning of V1 neurons. *Journal of Vision* **17**, 9–22 (2017).
40. Osaki, H., Naito, T., Sadakane, O., Okamoto, M. & Sato, H. Surround suppression by high spatial frequency stimuli in the cat primary visual cortex. *European Journal of Neuroscience* **33**, 923–932 (2011).
41. Serrano-Pedraza, I., Grady, J. P. & Read, J. C. A. Spatial frequency bandwidth of surround suppression tuning curves. *Journal of Vision* **12**, 24–24 (2012).
42. Kelly, D. H. Spatial frequency selectivity in the retina. *Vision Research* **15**, 665–672 (1975).
43. Skottun, B. C., Bradley, A. & Ramoa, A. S. Effect of contrast on spatial frequency tuning of neurones in area 17 of cat's visual cortex. *Experimental brain research* **63**, 431–435 (1986).
44. Webster, M. A. & Miyahara, E. Contrast adaptation and the spatial structure of natural images. *Journal of the Optical Society of America. A, Optics, image science, and vision* **14**, 2355–2366 (1997).
45. Turner, M. H., Schwartz, G. W. & Rieke, F. Receptive field center-surround interactions mediate context-dependent spatial contrast encoding in the retina. *bioRxiv* (2018).
46. Kuo, S. P., Schwartz, G. W. & Rieke, F. Nonlinear Spatiotemporal Integration by Electrical and Chemical Synapses in the Retina. *Neuron* **90**, 320–332 (2016).
47. Manookin, M. B., Patterson, S. S. & Linehan, C. M. Neural Mechanisms Mediating Motion Sensitivity in Parasol Ganglion Cells of the Primate Retina. *Neuron* **97**, 1–19 (2018).
48. Gollisch, T. Features and functions of nonlinear spatial integration by retinal ganglion cells. *Journal of Physiology-Paris* **111** (2013).
49. Butts, D. A. & Goldman, M. S. Tuning Curves, Neuronal Variability, and Sensory Coding. *PLoS Biology* **4**, e92 (2006).
50. Barlow, H. B. Possible principles underlying the transformation of sensory messages. *Sensory communication* (1961).
51. Barlow, H. Redundancy reduction revisited. *Network (Bristol, England)* **12**, 241–253 (2001).
52. Kayser, C. Processing of complex stimuli and natural scenes in the visual cortex. *Current opinion in neurobiology* **14**, 468–473 (2004).
53. Borst, A. & Theunissen, F. E. Information theory and neural coding. *Nature neuroscience* **2**, 947–957 (1999).
54. Botella-Soler, V., Deny, S., Martius, G., Marre, O. & Tkačik, G. Nonlinear decoding of a complex movie from the mammalian retina. *PLoS Computational Biology* **14**, e1006057–27 (2018).
55. Laughlin, S. B., de Ruyter van Steveninck, R. R. & Anderson, J. C. The metabolic cost of neural information. *Nature neuroscience* **1**, 36–41 (1998).
56. Chávez, A. E., Bozinovic, F., Peichl, L. & Palacios, A. G. Retinal Spectral Sensitivity, Fur Coloration, and Urine Reflectance in the Genus *Octodon* (Rodentia): Implications for Visual Ecology. *Investigative Ophthalmology & Visual Science* **44**, 2290–2296 (2003).
57. Litke, A. M. *et al.* What does the eye tell the brain? Development of a system for the large-scale recording of retinal output activity. *Nuclear Science, IEEE Transactions on* **51**, 1434–1440 (2004).
58. Segev, R., Goodhouse, J., Puchalla, J. & Berry, M. J. Recording spikes from a large fraction of the ganglion cells in a retinal patch. *Nature neuroscience* **7**, 1154–1161 (2004).
59. Yger, P. *et al.* Fast and accurate spike sorting *in vitro* and *in vivo* for up to thousands of electrodes. *Tech. Rep.* (2016).
60. Chichilnisky, E. J. A simple white noise analysis of neuronal light responses. *Network (Bristol, England)* **12**, 199–213 (2001).
61. Petrusca, D. *et al.* Identification and Characterization of a Y-Like Primate Retinal Ganglion Cell Type. *Journal of Neuroscience* **27**, 11019–11027 (2007).
62. Chichilnisky, E. J. & Kalmar, R. S. Functional asymmetries in ON and OFF ganglion cells of primate retina. *The Journal of neuroscience: the official journal of the Society for Neuroscience* **22**, 2737–2747 (2002).
63. Pinto, L. & Baron, J. Spatiotemporal Frequency Tuning Dynamics of Neurons in the Owl Visual Wulst. *Journal of Neurophysiology* **103**, 3424–3436 (2010).
64. Zylberberg, J. & DeWeese, M. R. Sparse Coding Models Can Exhibit Decreasing Sparseness while Learning Sparse Codes for Natural Images. *PLoS Computational Biology* **9**, e1003182 (2013).

Acknowledgements

Fondecyt 1150638 and 1140403, CONICYT-Basal Project FB0008, Millennium Institute ICM-P09-022-F. L.U.P. was supported by ANR projects “BalaV1” (ANR-13-BSV4-0014-02) and “TRAJECTORY” (ANR-15-CE37-0011). C.R.R. was supported by a CONICYT Ph.D. scholarship. We thank Dr. John Ewer for proofreading the manuscript.

Author Contributions

M.J.E., L.U.P., A.P. conceived the study. A.P. provided the animals, equipment for electrophysiology and data analysis. C.R.R. designed the stimulation protocols and conducted the experiments. C.R.R. and L.U.P. analyzed the results. C.R.R. and M.J.E. prepared the figures. All authors reviewed the manuscript.

Additional Information

Supplementary information accompanies this paper at <https://doi.org/10.1038/s41598-018-36861-8>.

Competing Interests: The authors declare no competing interests.

Publisher’s note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2019