Article

Current Biology

Consequences of the Oculomotor Cycle for the Dynamics of Perception

Highlights

- Eye movements transform spatial information into temporal modulations on the retina
- Saccades and fixational drift reformat space in time in complementary ways
- Oculomotor reformatting of visual input initiates coarse-to-fine processing
- Neurons in the retina and thalamus are tuned to oculomotor dynamics

Authors

Marco Boi, Martina Poletti, Jonathan D. Victor, Michele Rucci

Correspondence

mrucci@bu.edu

In Brief

Processing of large-scale visual information before fine spatial detail is believed to emerge from the response of neurons in the retina and thalamus. Using spectral analyses, neural modeling, and gaze-contingent control of retinal stimulation in humans, Boi et al. show that these coarse-to-fine dynamics are initiated by eye movements.

Boi et al., 2017, Current Biology *27*, 1–10
May 8, 2017
http://dx.doi.org/10.1016/j.cub.2017.03.034

CellPress

Please cite this article in press as: Boi et al., Consequences of the Oculomotor Cycle for the Dynamics of Perception, Current Biology (2017), http://dx.doi.org/10.1016/j.cub.2017.03.034

Current Biology
Article

Consequences of the Oculomotor Cycle for the Dynamics of Perception

Marco Boi,[1,4] Martina Poletti,[1,4] Jonathan D. Victor,[3] and Michele Rucci[1,2,5,*]
[1]Department of Psychological and Brain Sciences
[2]Graduate Program in Neuroscience
Boston University, 2 Cummington Mall, Boston, MA 02215, USA
[3]Feil Family Brain and Mind Research Institute, Weill Cornell Medical College, 1300 York Avenue, New York, NY 10065, USA
[4]These authors contributed equally
[5]Lead Contact
*Correspondence: mrucci@bu.edu
http://dx.doi.org/10.1016/j.cub.2017.03.034

SUMMARY

Much evidence indicates that humans and other species process large-scale visual information before fine spatial detail. Neurophysiological data obtained with paralyzed eyes suggest that this coarse-to-fine sequence results from spatiotemporal filtering by neurons in the early visual pathway. However, the eyes are normally never stationary: rapid gaze shifts (saccades) incessantly alternate with slow fixational movements. To investigate the consequences of this oculomotor cycle on the dynamics of perception, we combined spectral analysis of visual input signals, neural modeling, and gaze-contingent control of retinal stimulation in humans. We show that the saccade/fixation cycle reformats the flow impinging on the retina in a way that initiates coarse-to-fine processing at each fixation. This finding reveals that the visual system uses oculomotor-induced temporal modulations to sequentially encode different spatial components and suggests that, rather than initiating coarse-to-fine processing, spatiotemporal coupling in the early visual pathway builds on the information dynamics of the oculomotor cycle.

INTRODUCTION

A popular view of the early visual system compares the eye to a camera, where the camera lens represents the eye's optics and the film acts as the retina. Aside from obvious simplifications, there are two important aspects in which this analogy falls short. First, unlike the film in a camera, our retina and visual system are primarily sensitive to temporal modulations [1–5], to the point that even gross changes in the scene escape detection when they occur too slowly [6]. Therefore, while a camera needs a stationary image on its film to operate optimally, we see best when the input to our retina varies with time. Second, while a photographer usually takes great care to ensure that the camera does not move, our eyes are never at rest.

Two major types of eye movements continually alternate during examination of a static scene. Every few hundreds of milliseconds, saccades rapidly shift gaze toward new regions of interest [7]—sometimes even by very small amounts (microsaccades) [8, 9]—enabling examination of interesting stimuli with the foveola, the retinal region with highest acuity. In the "fixation" periods in between saccades, the eyes jitter incessantly following seemingly erratic trajectories that resemble Brownian motion [10, 11], a behavior often termed ocular drift [12]. Although not always considered in studies of visual functions, eye movements are always present and, thus, intertwined with the neural processing stages that encode visual information.

To date, most studies of this interaction focus on the challenges that eye movements pose to the establishment of visual representations [7, 13, 14]. In contrast, considerably less work has been dedicated to the analysis of a possible beneficial consequence of oculomotor activity, the reformatting of the visual input in the joint space-time domain [15, 16]. By modulating the signals experienced by retinal receptors, all eye movements transform the spatial pattern of luminance into a spatiotemporal flow on the retina. This transformation converts spatial variations into temporal changes, a format that neurons in the visual system greatly prefer [17, 18]. Because of their very different characteristics, saccades and drifts deliver signals with correspondingly different spatiotemporal statistics to the retina. Thus, the actual input to the visual system strongly depends on the observer's eye movements.

These considerations raise the interesting hypothesis that eye movements may enable selection of information not just in the spatial domain, by reorienting the fovea, but also in the temporal domain, by reformatting luminance signals so as to deliver selected aspects of spatial information within the temporal range of neuronal sensitivity. In previous research, we have shown that ocular drift yields a specific, beneficial frequency equalization during viewing of natural scenes [19]. However, little is known about how saccades reformat the retinal stimulus and about the effects of the normal alternation between rapid/large and slow/small eye movements. What are consequences of the saccade/fixation cycle for retinal stimulation? Do these stereotypical input changes affect perceptual dynamics? And, if so, what do these effects imply for the encoding of spatial information? Here we show that this oculomotor cycle plays a key functional role: it initiates a coarse-to-fine processing sequence at each visual fixation.

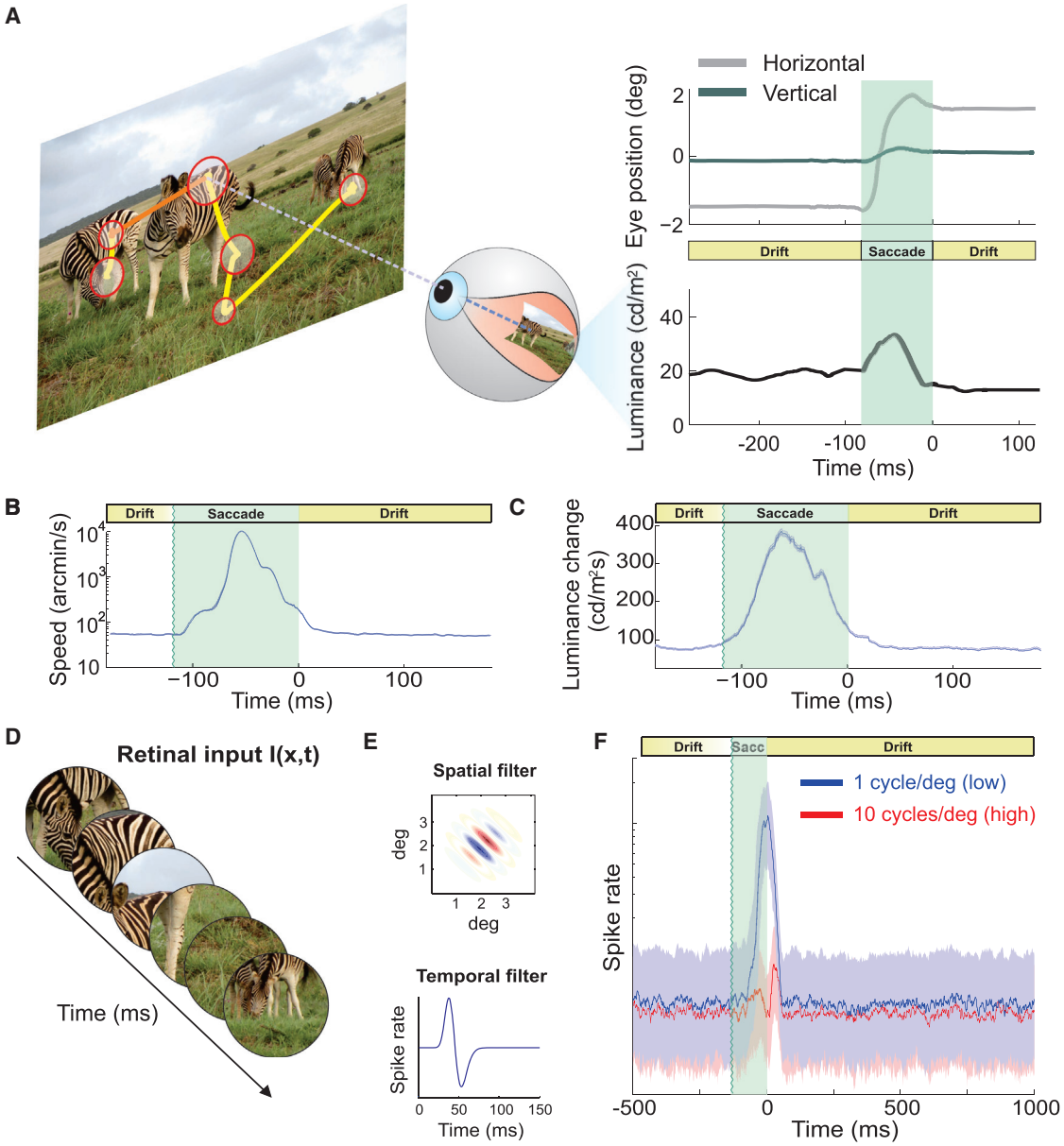

Figure 1. Natural Input to the Retina and Predicted Neural Dynamics

(A) Normal alternation between saccades (yellow segments) and periods of fixational eye movements (red circles) during viewing of natural scenes (image size: 17° × 12.8°). The size of each circle represents fixation duration. A portion of the eye movement trace (orange) is expanded in the right panel (top), together with the luminance signal experienced by a retinal receptor (bottom). Eye movements continually modulate the input flow onto the retina, effectively redistributing spatial information in the joint space-time domain.

(B and C) Mean instantaneous speed (B) of the retinal image (note log scale) and mean rate of change in the input luminance (C) experienced by a foveal receptor during a fixation-saccade-fixation sequence. Time zero marks saccade end.

(D–F) Neural modeling. (D) The responses of model V1 simple cells were simulated as their receptive fields moved following recorded oculomotor traces. (E) Models consisted of rectified filters with separable spatial (top) and temporal (bottom) kernels. (F) Responses of neurons tuned to low (1 cycle/degree, blue line) and high spatial frequency (10 cycles/degree, red line). Solid lines represent medians of activity. Shaded regions enclose the range between the first and third quartiles.

RESULTS

Retinal Input Dynamics and Modeling Predictions

We started by recording oculomotor activity while human observers freely examined pictures of natural scenes (Figure 1A). As is typical, subjects frequently relocated their gaze by means of saccades (average inter-saccadic duration ± SD: 264 ± 78 ms), which spanned a wide range of amplitudes, from a fraction of a degree to several degrees (mean amplitude: 4.4° ± 1.3°), and reached velocities as high as 900°/s. As

expected from previous studies [7, 12, 20], the eye remained in motion in the periods in between saccades, when it drifted with apparently random trajectories at an average speed of almost 1°/s (Figure 1B). Thus, eye movements continually transformed spatial information into temporal modulations on the retina, yielding motion signals that would be immediately visible had they originated from the motion of objects in the scene rather than the observer's eye movements. This recurrent alternation between macroscopic and microscopic eye movements results in a stereotypical sequence of retinal stimulation, with the brief abrupt changes caused by saccades at fixation onset followed by the slower luminance changes caused by ocular drifts as fixation progresses (Figure 1C).

We examined the impact of this time-varying input on the responses of modeled simple cells in the primary visual cortex (V1). Models were exposed to the same visual input experienced by the subjects in our experiments, i.e., the spatiotemporal luminance signals impinging onto the retinas given the recorded sequences of eye movements (Figures 1D and 1E). We simulated the mean instantaneous firing rates of neurons with peak sensitivity at either low (1 cycle/degree) or high (10 cycles/degree) spatial frequencies. Apart from the difference in spatial frequency tuning, these two neuronal populations were identical; yet, they responded very differently during the course of post-saccadic fixation (Figure 1F). At fixation onset, the temporal transient caused by the preceding saccade elicited a strong response in units tuned to low spatial frequency, but had a modest impact on high-frequency neurons. In contrast, later in fixation, once the effect of the preceding saccade had terminated, the two sets of neurons gave comparable responses, even though the images presented to the observer contained ~100 times more power at low than high spatial frequencies. Thus, modeled neural responses transiently followed the power distribution of natural images immediately after a saccade, and then they settled into a more balanced regime with equal responses in different frequency channels during the course of fixation.

To understand the factors responsible for these neural dynamics, we examined the frequency content of the visual input transients resulting from different types of eye movements. When a static scene was observed in the complete absence of oculomotor activity, the input to the retina was a static image, so its power was confined to 0 Hz, the temporal DC frequency axis (Figure 2A). Eye movements transformed this static input into a spatiotemporal stimulus on the retina, redistributing its power across nonzero temporal frequencies in a way that depends on the specific characteristics of eye movements (Figure 2B). Since drifts and saccades move the eyes in very different ways, one may expect considerably different power redistributions driving neural responses during the early and late phases of post-saccadic fixation.

We first examined the spatiotemporal power redistribution present during late fixation, when input signals were only modulated by ocular drift (Figures 2C and 2D). We have recently shown that the amount of DC power that drifts convert into nonzero temporal frequencies increases with spatial frequency in a way that counterbalances the spectral density of natural scenes [19], effectively a form of matching between the characteristics of eye movements and those of the natural world. Because of this interplay, ocular drift produces a striking spatiotemporal reformatting of the input to the retina during viewing of natural scenes. Unlike the power-law spectral density of the images displayed on the monitor [21] (black line in Figure 2D), the temporal modulations resulting from ocular drift contained equal power over a broad range of spatial frequencies (Figures 2C and 2D), an effect known as spectral whitening.

The abrupt input changes caused by saccades possessed very different spectral densities (Figures 2E and 2F). Unlike drift, saccades converted spatial power into temporal power at a similar rate across a broad range of spatial frequencies. Therefore, saccade transients preserved the uneven spectral distributions of the scenes presented to the observers: at all relevant nonzero temporal frequencies, the power of the visual input to the retina declined with the square of spatial frequency, like the spectrum of natural images. Because of this effect, the total temporal power (the sum across all nonzero temporal frequencies) transiently delivered by saccades was larger than that resulting from ocular drift at low spatial frequencies, and the two became comparable around 10 cycles/degree.

The spectral redistributions in Figure 2 explain the dynamics of neural activity in our model (Figure 1F). They show that the normal alternation between saccades and fixational drift sequentially exposes neurons to modulations that contain different kinds of spatial information during the course of normal post-saccadic fixation. At fixation onset, because of the integration memory of the neuronal firing (the temporal sensitivity profile in Figure 1E), neurons are effectively stimulated by an input signal dominated by low spatial frequencies, as it occurs in natural images. Later during fixation, ocular drift enhances high spatial frequencies, and the same neurons are now driven by a signal in which power has been equalized across spatial frequencies.

To make quantitative predictions of the perceptual consequences of these retinal dynamics, we modeled the post-saccadic course of contrast sensitivity in a standard detection experiment. Our modeled V1 neurons were now exposed to a grating embedded within a natural noise field, as their receptive fields translated following recorded sequences of eye movements (Figure 3A). A standard decision-making model continually accumulated the evidence provided by neural activity during post-saccadic fixation and determined contrast thresholds by comparing the mean instantaneous firing rates simulated in the presence and absence of the grating (Figure 3B).

In keeping with the spectral analysis in Figure 2, saccades and ocular drift yielded different modulations during exposure to noisy gratings at 1 and 10 cycles/degree. These modulations effectively varied the signal-to-noise Ratio (SNR) in the input to the decision-making stage, as they changed the level of stimulus-driven response relative to the fixed internal noise of the model. Responses to a 1 cycle/degree grating were clearly distinct from those generated by the noise field alone only at fixation onset (early fixation; Figure 3C, top). In contrast, because of the high-frequency enhancement given by ocular drift, neural responses at 10 cycles/degree remained distinguishable from noise during both early and late fixation (Figure 3C, bottom). In fact, for these neurons, the separation between the responses with and without the target was comparably large in early and late fixation. Thus, at 1 cycle/degree, detection was greatly facilitated immediately after a saccade, and it did not benefit from

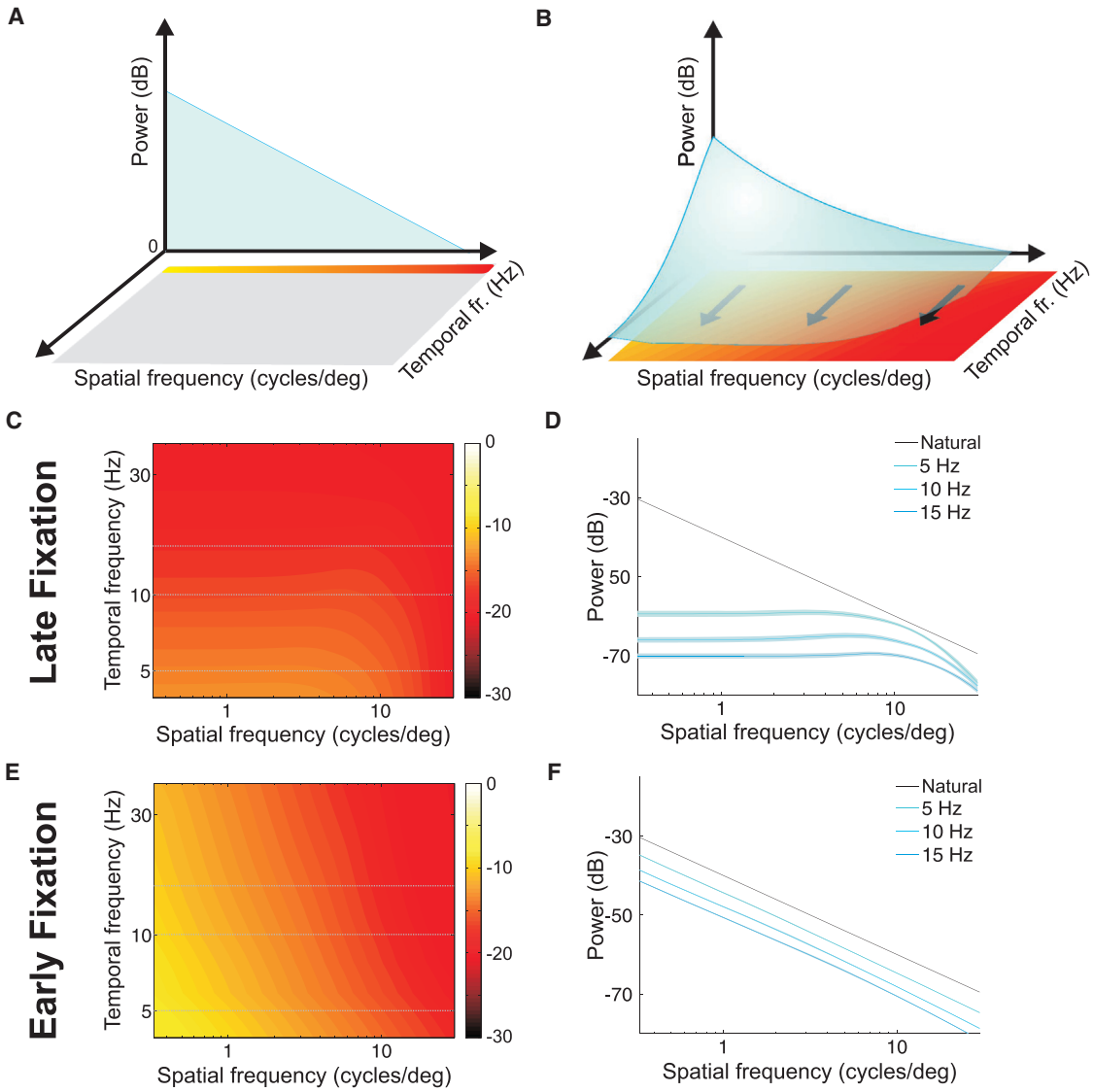

Figure 2. Power Spectra of the Retinal Input Signals Resulting from Different Types of Eye Movements

(A and B) Eye movements transform spatial information in the scene into spatiotemporal modulations on the retina, redistributing the power at 0 Hz of a static scene (A) across non-zero temporal frequencies (B). The characteristics of this transformation depend on the type of eye movements.

(C and D) Spectral density of the visual input present during fixational drift. Both the full spatiotemporal distribution (C) and sections at individual temporal frequencies (D) are shown. Sections in (D) are compared to the ideal $1/k^2$ power spectrum of natural images (Natural). During viewing of natural scenes, fixational drift equalizes power across a broad range of spatial frequencies.

(E and F) Spectral density of the visual input given by saccades. Unlike ocular drift, saccades yield modulations that follow the spectral density of natural images over a wide range of temporal frequencies. Both the full spatiotemporal distribution (E) and sections at individual temporal frequencies (F) are shown.

integration of firing rates past early fixation, whereas at 10 cycles/degree, continued temporal integration was beneficial, as it further separated the response distributions elicited in the presence and absence of the grating (Figure 3D). Because of these effects, modeled sensitivity remained virtually constant at low spatial frequencies during the course of fixation, but it continued to increase at high spatial frequencies (Figure 3E).

These results suggest distinct perceptual contributions from different types of eye movements. Three specific predictions emerge as follows: (1) the transients introduced by saccades are primarily responsible for vision at low spatial frequencies, a process that occurs rapidly following fixation onset; (2) the modulations resulting from ocular drift continually enhance high spatial frequencies at later times during post-saccadic fixation; and (3) more broadly, the normal alternation between saccades and fixational drift gives rise to a stereotypical time course of visual perception, leading to an *oculomotor-induced* coarse-to-fine analysis of the scene.

Experimental Validation of Theoretical Predictions

We tested our predictions by assessing the consequences of the normal saccade/drift cycle on visual sensitivity in human

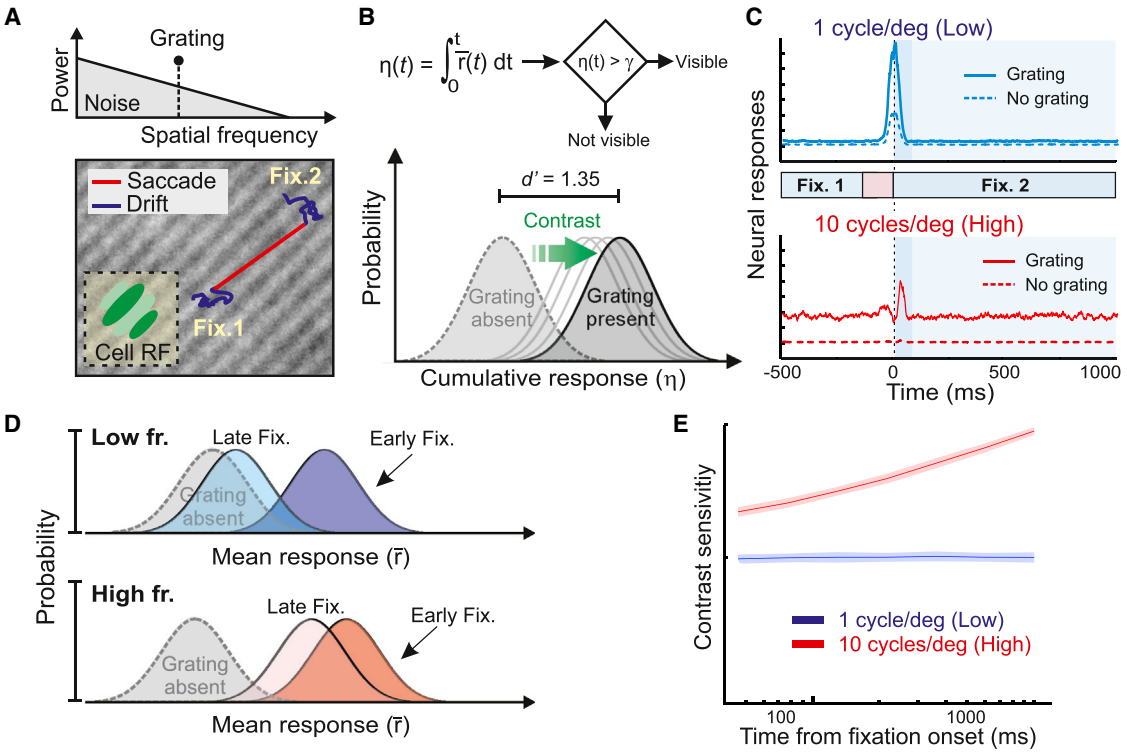

Figure 3. Predicted Dynamics of Contrast Sensitivity

(A) The responses of V1 neurons were modeled in a detection task with gratings embedded in a naturalistic noise field (top). Models were tuned to the grating's orientation and spatial frequency (either 1 or 10 cycles/degree), and their receptive fields were translated following sequences of recorded eye movements (two fixations separated by a saccade; bottom).

(B) At each time during post-saccadic fixation, a standard decision-making model determined the presence or absence of the grating based on the cumulative neural response (averaged across all simulated cells) from fixation onset ($\eta(t)$; top). The contrast yielding a hit rate of 0.75 and false alarm rate of 0.25 ($d' = 1.35$) was selected as the threshold (bottom).

(C) Average responses with or without the grating at 1 (top) and 10 cycles/degree (bottom). Shaded regions mark the post-saccadic period of fixation considered by the model in (B). The darker shade represents the period of early fixation in which the preceding saccade influences neural responses. Note how responses to the high spatial frequency target remain well separated from the noise response not only in early fixation but also in late fixation.

(D) Schematic representation of the impact of different types of eye movements on visual detection. For any given contrast, the separation between the distributions of cell responses in the presence and absence of the grating transiently increases following saccades at low spatial frequencies but remains more similar throughout the course of fixation at high spatial frequencies.

(E) Model predictions. With a 1 cycle/degree grating, contrast sensitivity does not increase beyond the early-fixation level provided by the saccade transient. Sensitivity to 10 cycles/degree improves more gradually, as it also weights the amplified response that persists during late fixation because of ocular drift.

observers. We first determined whether, following a saccade, spatial vision evolves as predicted by Figure 3E. In two further experiments, we then separately analyzed the contributions of ocular drift and saccades by selectively eliminating their input modulations.

Although eye movements are known to both suppress and enhance visual functions [22, 23], standard laboratory measurements of contrast sensitivity typically ignore or attempt to minimize the effects of the luminance changes resulting from oculomotor activity. As a consequence, little is known about visual dynamics during normal inter-saccadic fixation. As in the simulations of Figure 3, we measured the post-saccadic evolution of contrast thresholds at low and high spatial frequencies (1 and 10 cycles/degree, respectively; Figure 4C). To maintain control of the spatiotemporal stimulus impinging onto the retina while ensuring foveal exposure to the normal modulations resulting from the saccade/fixation cycle, we synchronized stimulus delivery to eye movements and tested sensitivity at different times (100 and 800 ms) after a saccade had swept the retina across a naturalistic noise field (Figures 4A and 4B). In keeping with our predictions, sensitivity to low spatial frequencies changed little during the course of fixation: after 800 ms exposure, thresholds measured at 1 cycle/degree were similar to those measured after only 100 ms ($p > 0.24$, paired two-tailed t test; Figure 4D). In contrast, prolonged stimulus exposure was beneficial at high spatial frequencies, leading to a >2-fold elevation of sensitivity at 10 cycles/degree ($p < 0.027$), an effect present in each individual observer ($p \leq 0.043$, parametric bootstrap [24]). Thus, performance improved over the course of fixation at high, but not low, spatial frequencies. These results confirm the general prediction of our model: visual perception follows coarse-to-fine dynamics during normal post-saccadic fixation over a naturalistic field.

Having established the presence of the predicted post-saccadic dynamics of sensitivity, we examined whether these dynamics were driven by the luminance modulations caused by eye movements, as predicated by our theoretical analyses.

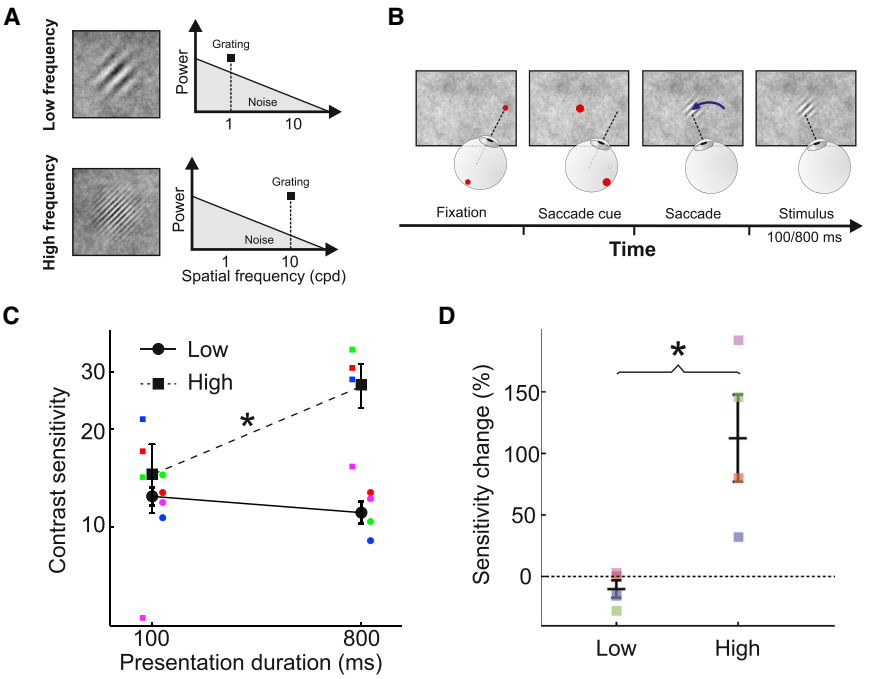

Figure 4. Post-saccadic Dynamics of Contrast Sensitivity

(A and B) Experimental procedure. (A) Subjects reported the orientation (±45°) of a grating at either 1 (low frequency; top) or 10 cycles/degree (high frequency; bottom) embedded in a full-screen naturalistic noise field ($1/k^2$ spectrum; shaded triangle). (B) Subjects initially fixated ~7° away and performed a saccade toward the future grating location at the appearance of a cue. The stimulus appeared during the saccade and was displayed for either 100 or 800 ms following fixation onset.

(C) Contrast sensitivity (n = 4). Sensitivity increased with longer stimulus exposure for high, but not low, spatial frequencies.

(D) Relative change in sensitivity with exposure duration. In both (C) and (D), colored symbols are individual subject data, black symbols are group means, and error bars represent SEM. Asterisks mark statistically significant differences ($p < 0.03$, paired two-tailed t test).

To specifically test the impact of ocular drift and saccades at individual spatial frequencies, we repeated the experiment of Figure 4 with the presentation of isolated gratings and manipulated the oculomotor modulations of retinal input signals by means of gaze-contingent display control procedures.

To isolate the influence of ocular drift, we compared contrast thresholds measured at late fixation (800 ms post-saccadic exposure) in the presence and absence of retinal image motion. It is known that high-acuity discriminations are impaired under retinal stabilization [25, 26]. These previous studies, however, did not separate the possible contributions of different types of fixational eye movements—ocular drift and small saccades (microsaccades)—and it has been recently demonstrated that microsaccade centering of the stimulus within the foveola facilitates fine spatial vision [27]. Here we discarded all trials containing microsaccades and specifically focused on the impact of the luminance modulations caused by ocular drift. To selectively eliminate drift fluctuations while preserving the initial saccadic transient, we stabilized the stimulus on the observer's retina at the detection of saccade end. That is, following the initial saccade, we continually adjusted the position of the stimulus on the display according to the subject's eye movements, so that it remained at a fixed position on the retina [26, 28].

Retinal stabilization acted differently on the time course of sensitivity at high and low spatial frequencies (Figure 5B). As predicted, drift cancellation had minimal impact at low spatial frequencies: thresholds measured at 1 cycle/degree were very similar and statistically indistinguishable during normal viewing and under retinal stabilization ($p > 0.43$). In contrast, removal of retinal image motion greatly affected sensitivity at high spatial frequency, attenuating the improvement measured during the course of normal unstabilized fixation by about 40% in both observers. These results confirm our modeling prediction that continual post-saccadic integration of drift luminance modulations selectively enhances high-frequency sensitivity as fixation progresses.

To isolate the influences of the luminance transients caused by saccades, we relied on gaze-contingent control of stimulus contrast. As in the previous experiments, subjects performed a saccade toward a cued location. Here, however, the abrupt change in the retinal input resulting from the saccade (normal condition) was replaced by a gradual increment in the contrast of the grating timed with saccade landing (no-transient condition). That is, the 800 ms period of fixed-contrast stimulus exposure in the experiments of Figures 4 and 5 was here preceded by a 1.5 s contrast ramp starting at fixation onset (Figure 6A). Thus, in the no-transient condition, subjects were exposed to the stimulus for a much longer time, receiving almost double the energy experienced in the normal condition, but were deprived of saccadic transients.

As predicted, the consequences of saccades were complementary to those of fixational drift (Figure 6B): elimination of the saccade transient impaired sensitivity in each individual observer ($p \leq 0.0035$, parametric bootstrap). On average across observers, removal of the saccade transient lowered sensitivity by ~40% at 1 cycle/degree ($p = 0.017$, two-tailed paired t test), but it had little impact at 10 cycles/degree ($p = 0.93$), resulting in a much larger change in sensitivity at low spatial frequency than high spatial frequency ($p = 0.047$, two-tailed paired t test). In other words, at low spatial frequencies, saccades cannot be traded for ocular drifts by increasing the duration of fixation, as drift modulations are not beneficial in this range regardless of their duration. These findings confirm our theoretical prediction that saccade transients are critical for low spatial frequency vision.

DISCUSSION

Species with a high-acuity fovea, like humans and other primates, redirect their gaze to examine objects of interest, but

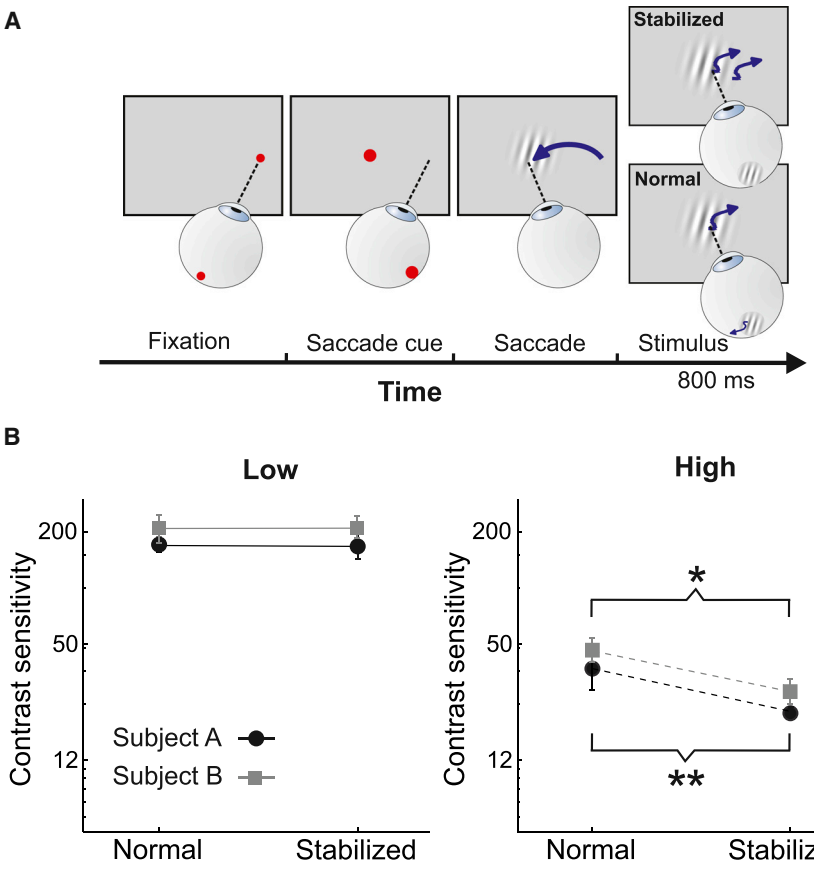

Figure 5. Perceptual Consequence of Fixational Drift

(A) Experimental procedure. Following the initial saccade, stimuli were either displayed at a fixed location on the monitor (Normal) or at a fixed location on the retina (Stabilized). In this latter condition, stimuli moved on the screen under real-time control to compensate for the subject's eye movements.

(B) Experimental results. Contrast sensitivity in the two viewing conditions at 1 (low) and 10 cycles/degree (high). Eliminating drift modulations impairs sensitivity at high (left), but not low (right), spatial frequencies. Each graph shows data from two observers. Error bars represent 95% confidence intervals. *p = 0.0035 and **p = 0.0005, parametric bootstrap test.

they also continue to move their eyes during the post-saccadic periods in which visual information is acquired and processed. Our results show that this behavioral cycle converts spatial information into temporal modulations in a way that generates coarse-to-fine dynamics of visual analysis within each individual fixation. This transformation bears fundamental implications for the way space is represented in the visual system, and it can be regarded as a primary function of oculomotor activity. Eye movements enable selection of visual information not just in the spatial domain, by properly redirecting the fovea, but also in the temporal domain, by determining what aspects of the visual scene fall within the temporal range of neuronal sensitivity.

The evidence presented in this study comes from three distinct components. First, spectral analyses of signals impinging onto retinal receptors showed that saccades and drift redistribute the power of the external stimulus in complementary ways on the retina. Second, modeling of V1 neurons predicted different consequences of these two regimes on visibility, with low spatial frequencies emerging rapidly after a saccade and high spatial frequencies benefiting from continued integration during the course of fixation. Third, psychophysical experiments validated these predictions by controlling retinal stimulation during the saccade/fixation cycle in human subjects. We found that, as predicted by our model, the evolution of contrast sensitivity during natural post-saccadic fixation exhibits *oculomotor-dependent* phasic and tonic dynamics at low and high spatial frequencies, respectively. Note that, in everyday viewing, these dynamics occur at each fixation and follow the reduction in visibility and the spatial updating phenomena present before and during saccades [13, 29].

A large body of evidence supports a coarse-to-fine hierarchy of visual analysis [30–34]. However, little consideration has been given to the temporal characteristics of the retinal stimulus itself, which results from the interaction of largely static visual scenes with the dynamics of eye movements. Our results indicate that this interaction is a critical component in establishing the coarse-to-fine processing sequence. In the laboratory, where observers are often required to maintain fixation, sharp stimulus onsets commonly replace saccades. Under natural viewing conditions, however, the saccade/fixation cycle is the primary source of temporal modulations. Note that because of the impossibility of completely eliminating oculomotor transients, both sets of data in Figures 5 and 6 most likely underestimate the real effects of eye movements. This happens because also the residual temporal modulations caused by imperfections in retinal stabilization in Figure 5 and by the contrast ramp in Figure 6 are expected to emphasize high and low spatial frequencies, respectively. Similarly in Figure 4, while our experiments provided natural stimulation in the fovea, the sharp peripheral onset of the stimulus during the saccade might have enhanced performance. Thus, eye movements may normally exert even stronger influences than those revealed by our data.

These findings provide critical insights into the strategy followed by the visual system to represent space. While the *spatial format* of neural representations (e.g., retinotopic versus spatiotopic [14, 35]) has been intensely debated, considerable less attention has been dedicated to the *multidimensional reformatting* of visual input signals given by eye movements, i.e., the conversion of spatial patterns into temporal modulations on the retina [15, 16]. Our findings show that the visual system takes advantage of the stereotypical evolution in the spectral redistribution of the input flow to sequentially represent space in the temporal domain during natural post-saccadic fixation. That is, spatial information is not only encoded by the position of neurons

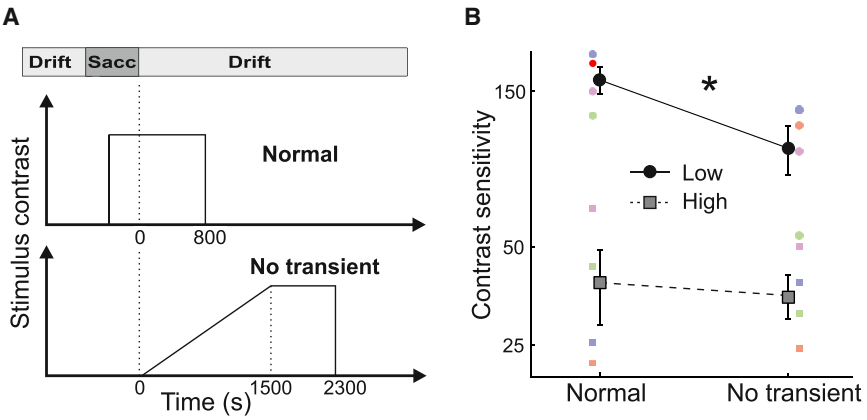

Figure 6. Perceptual Consequences of Saccade Transients

(A) Experimental procedure. The stimulus was either displayed during the saccade (Normal condition) or gradually increased in contrast after saccade end (No-transient condition). The graphs show the dynamics of stimulus contrast in the two conditions.

(B) Experimental results. Contrast sensitivity in the presence and absence of saccade transients for gratings at 1 and 10 cycles/degree. Small colored symbols are individual subject data, and large gray and black symbols are group means. Error bars represent SEM; *$p < 0.02$, paired two-tailed t test; $n = 4$.

within maps but also by the temporal structure of neural responses, dynamics governed by oculomotor activity and, in large part, under the observer's control. Saccades and drifts can be regarded as two fundamental computational elements of this strategy, as they deliver useful temporal modulations in complementary spatial frequency ranges.

This view implies important mechanisms for decoding visual signals beyond those commonly postulated. Since neural firing in thalamic afferents carries distinct kinds of information during early and late fixation, the input to the cortex needs to be interpreted according to the time elapsed from fixation onset. Neural responses immediately following a saccade are primarily driven by low spatial frequencies and presumably contribute to perception of the general structure of the scene. Later responses from the same neurons are in contrast driven by a whitened input and most likely signal luminance discontinuities and finer scales. These considerations support a fundamental role for the magnocellular pathway in the representation of the gist of a scene [36, 37] via its saccade-induced response, establishing a scaffolding for the later addition of fine spatial detail during fixational eye movements.

Our results also provide a functional explanation for the different dynamics of neurons sensitive to low and high spatial frequencies in the early visual system. It is well established that magnocellular and parvocellular neurons in the retina and lateral geniculate nucleus tend to prefer separate spatial frequency ranges, with magnocellular neurons sensitive to lower frequencies than parvocellular ones [38]. Notably, these neurons also exhibit different temporal preferences: magnocellular neurons tend to be more responsive to temporal transients than parvocellular neurons [39]. But why should neurons sensitive to separate spatial frequency ranges also differ in their temporal characteristics? The results in Figure 3 suggest an answer to these questions. In a system designed to encode space in time—i.e., to operate at non-zero temporal frequencies—analysis of low, but not high, spatial frequencies tends to be more reliable immediately following a saccade than later during fixation. In other words, an observer that optimally accrues spatial information from the oculomotor-shaped retinal input flow should (1) heavily rely on saccade transients to extract low spatial frequencies and (2) continue integrating throughout the course of fixation to extract high spatial frequencies. Thus, the incessant presence of eye movements implies distinct strategies for efficiently acquiring information in different frequency ranges, and the spatiotemporal characteristics of magnocellular and parvocellular neurons are well suited to carry out these functions.

The previous observations suggest that several neurophysiological findings with anesthetized and paralyzed animals need to be reevaluated in light of their implications for natural viewing conditions. Various studies have observed temporal changes in the responses of cortical neurons consistent with a progressive refinement of visual sensitivity, such as a sharp increase in preferred spatial frequency after stimulus presentation [40–42]. These cortical changes, which most likely originate from delays in thalamic afferents [43], are believed to represent the neural mechanisms responsible for coarse-to-fine perceptual dynamics. However, the view emerging from our study is that these responses constitute a direct consequence of a system optimally tuned to its input, once eye movements are taken into account: spatiotemporally coupled neural dynamics are not the primary cause for a coarse-to-fine evolution, but they are necessary to operate synergistically with oculomotor activity and efficiently extract spatial information from oculomotor transients. We predict that a coarse-to-fine perceptual course would still occur even in the absence of these neural delays as a consequence of the normal saccade/drift cycle, as demonstrated by our model in which all simulated neurons possessed identical temporal profiles.

Studies of visual perception often focus on the costs associated with gaze shifts, and how the visual system avoids their potentially disruptive consequences. Here we show that these gaze shifts have *computational advantages*. We have shown that the incessant alternation between macroscopic and microscopic eye movements leads to a dynamic reformatting of the visual flow, which enables encoding of space in the temporal domain. This space-time conversion acts as a first processing stage that operates together with the response selectivity of retinal neurons to implement a hierarchical scheme of visual analysis. An interesting hypothesis now emerges. In principle, control over eye movements could go beyond foveation to select the scale of visual processing and emphasize task-relevant ranges of spatial frequencies. Further work is needed to test whether humans and other species employ this form of active control.

EXPERIMENTAL PROCEDURES

Stimuli consisted of natural scenes (free-viewing experiment, Figures 1 and 2) and Gabor patches (forced-choice experiments, Figures 4, 5, and 6). They were displayed on a gamma-corrected fast-phosphor CRT monitor (Iiyama HM204DT) at either 100 Hz (free viewing) or 200 Hz (forced choice). Stimuli were observed monocularly with the left eye patched; 21 emmetropic subjects participated in the study, following the procedures approved by the Boston University Charles River Campus Institutional Review Board.

Eye movements were recorded by means of a Dual Purkinje Image eye-tracker (Fourward Technology) and sampled at 1 kHz, while a custom dental-imprint bite bar and a headrest prevented head movements. Movements with speeds >3°/s and amplitudes exceeding 3 arcmin were classified as saccades. Saccade amplitude was defined as the length of the vector connecting the two locations at which velocity became greater (onset) and lower (offset) than 2°/s.

Thresholds in Figures 4, 5, and 6 represent Michelson contrasts yielding 75% performance. The stimulus contrast changed at each trial following the PEST (parameter estimation by sequential testing) algorithm [44]. Gaze-contingent manipulation of the retinal input was achieved by means of a custom-developed system for flexible gaze-contingent display control [45], an extensively tested system that enables precise synchronization between eye movements and changes in the display. All trials in which the saccade landed more than 1° away from the center of stimulus or in which blinks or microsaccades occurred during stimulus presentation were discarded (see Supplemental Experimental Procedures for details).

Frequency analyses in Figure 2 were conducted by means of the eye movement displacement probability, the probability that the eyes moves by x in the interval t, an approach described in previous publications [19, 46] that enables high-resolution spectral estimation. The mean instantaneous firing rates of simple cells were modeled by means of standard space-time separable filters designed on the basis of neurophysiological data (Figure 1E). Modeled receptive fields moved following recorded traces of eye movements (Figures 1E and 3A). In Figure 3, a decision-making stage continually integrated the responses of modeled neurons tuned to the grating's frequency and orientation. At each time during post-saccadic fixation, contrast thresholds were selected as the contrast values yielding hits and false alarm rates of 0.75 and 0.25, respectively ($d' = 1.35$).

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures and can be found with this article online at http://dx.doi.org/10.1016/j.cub.2017.03.034.

AUTHOR CONTRIBUTIONS

M.B. implemented the experiments. M.B. and M.P. collected and analyzed data. J.D.V. and M.R. supervised the project. All authors contributed to the design of the study and the writing of the manuscript.

ACKNOWLEDGMENTS

This work was supported by NIH grants EY18363 to M.R. and EY07977 to J.V. and National Science Foundation grants 1457238, 1420212 to M.R., and BCS-1534932 to M.P.

Received: February 3, 2017
Revised: March 11, 2017
Accepted: March 15, 2017
Published: April 20, 2017

REFERENCES

1. Robson, J.G. (1966). Spatial and temporal contrast-sensitivity functions of the visual system. J. Opt. Soc. Am. *56*, 1141–1142.
2. Nagano, T. (1980). Temporal sensitivity of the human visual system to sinusoidal gratings. J. Opt. Soc. Am. *70*, 711–716.
3. Breitmeyer, B., and Julesz, B. (1975). The role of on and off transients in determining the psychophysical spatial frequency response. Vision Res. *15*, 411–415.
4. Benardete, E.A., and Kaplan, E. (1999). The dynamics of primate M retinal ganglion cells. Vis. Neurosci. *16*, 355–368.
5. Benardete, E.A., and Kaplan, E. (1997). The receptive field of the primate P retinal ganglion cell, I: Linear dynamics. Vis. Neurosci. *14*, 169–185.
6. Rensink, R.A. (2002). Change detection. Annu. Rev. Psychol. *53*, 245–277.
7. Kowler, E. (2011). Eye movements: the past 25 years. Vision Res. *51*, 1457–1483.
8. Collewijn, H., and Kowler, E. (2008). The significance of microsaccades for vision and oculomotor control. J. Vis. *8*, 1–21.
9. Poletti, M., and Rucci, M. (2016). A compact field guide to the study of microsaccades: challenges and functions. Vision Res. *118*, 83–97.
10. Ratliff, F., and Riggs, L.A. (1950). Involuntary motions of the eye during monocular fixation. J. Exp. Psychol. *40*, 687–701.
11. Steinman, R.M., Haddad, G.M., Skavenski, A.A., and Wyman, D. (1973). Miniature eye movement. Science *181*, 810–819.
12. Rucci, M., and Poletti, M. (2015). Control and functions of fixational eye movements. Annu. Rev. Vis. Sci. *1*, 499–518.
13. Wurtz, R.H. (2008). Neuronal mechanisms of visual stability. Vision Res. *48*, 2070–2089.
14. Burr, D.C., and Morrone, M.C. (2011). Spatiotopic coding and remapping in humans. Philos. Trans. R. Soc. Lond. B Biol. Sci. *366*, 504–515.
15. Ahissar, E., and Arieli, A. (2001). Figuring space by time. Neuron *32*, 185–201.
16. Rucci, M., and Victor, J.D. (2015). The unsteady eye: an information-processing stage, not a bug. Trends Neurosci. *38*, 195–206.
17. Croner, L.J., and Kaplan, E. (1995). Receptive fields of P and M ganglion cells across the primate retina. Vision Res. *35*, 7–24.
18. Kagan, I., Gur, M., and Snodderly, D.M. (2008). Saccades and drifts differentially modulate neuronal activity in V1: effects of retinal image motion, position, and extraretinal influences. J. Vis. *8*, 1–25.
19. Kuang, X., Poletti, M., Victor, J.D., and Rucci, M. (2012). Temporal encoding of spatial information during active visual fixation. Curr. Biol. *22*, 510–514.
20. Poletti, M., Aytekin, M., and Rucci, M. (2015). Head-eye coordination at a microscopic scale. Curr. Biol. *25*, 3253–3259.
21. Field, D.J. (1987). Relations between the statistics of natural images and the response properties of cortical cells. J. Opt. Soc. Am. A *4*, 2379–2394.
22. Ross, J., Burr, D., and Morrone, C. (1996). Suppression of the magnocellular pathway during saccades. Behav. Brain Res. *80*, 1–8.
23. Ibbotson, M., and Krekelberg, B. (2011). Visual perception and saccadic eye movements. Curr. Opin. Neurobiol. *21*, 553–558.
24. Wichmann, F.A., and Hill, N.J. (2001). The psychometric function: II. Bootstrap-based confidence intervals and sampling. Percept. Psychophys. *63*, 1314–1329.
25. Rucci, M., and Desbordes, G. (2003). Contributions of fixational eye movements to the discrimination of briefly presented stimuli. J. Vis. *3*, 852–864.
26. Rucci, M., Iovin, R., Poletti, M., and Santini, F. (2007). Miniature eye movements enhance fine spatial detail. Nature *447*, 851–854.
27. Poletti, M., Listorti, C., and Rucci, M. (2013). Microscopic eye movements compensate for nonhomogeneous vision within the fovea. Curr. Biol. *23*, 1691–1695.
28. Poletti, M., and Rucci, M. (2010). Eye movements under various conditions of image fading. J. Vis. *10*, 1–18.
29. Ross, J., Morrone, M.C., Goldberg, M.E., and Burr, D.C. (2001). Changes in visual perception at the time of saccades. Trends Neurosci. *24*, 113–121.

30. Burr, D.C. (1981). Temporal summation of moving images by the human visual system. Proc. R. Soc. Lond. B Biol. Sci. *211*, 321–339.
31. Watt, R.J. (1987). Scanning from coarse to fine spatial scales in the human visual system after the onset of a stimulus. J. Opt. Soc. Am. A *4*, 2006–2021.
32. Schyns, P.G., and Oliva, A. (1994). From blobs to boundary edges: evidence for time- and spatial-scale-dependent scene recognition. Psychol. Sci. *5*, 195–200.
33. Hegdé, J. (2008). Time course of visual perception: coarse-to-fine processing and beyond. Prog. Neurobiol. *84*, 405–439.
34. Neri, P. (2011). Coarse to fine dynamics of monocular and binocular processing in human pattern vision. Proc. Natl. Acad. Sci. USA *108*, 10726–10731.
35. Hall, N.J., and Colby, C.L. (2011). Remapping for visual stability. Philos. Trans. R. Soc. Lond. B Biol. Sci. *366*, 528–539.
36. Bar, M., Kassam, K.S., Ghuman, A.S., Boshyan, J., Schmid, A.M., Dale, A.M., Hämäläinen, M.S., Marinkovic, K., Schacter, D.L., Rosen, B.R., and Halgren, E. (2006). Top-down facilitation of visual recognition. Proc. Natl. Acad. Sci. USA *103*, 449–454.
37. Kveraga, K., Boshyan, J., and Bar, M. (2007). Magnocellular projections as the trigger of top-down facilitation in recognition. J. Neurosci. *27*, 13232–13240.
38. Schiller, P.H., and Logothetis, N.K. (1990). The color-opponent and broad-band channels of the primate visual system. Trends Neurosci. *13*, 392–398.
39. Casagrande, V.A., and Xu, X. (2004). Parallel visual pathways: a comparative perspective. In The Visual Neurosciences, L.M. Chalupa, and J.S. Werner, eds. (The MIT Press), pp. 494–506.
40. Bredfeldt, C.E., and Ringach, D.L. (2002). Dynamics of spatial frequency tuning in macaque V1. J. Neurosci. *22*, 1976–1984.
41. Malone, B.J., Kumar, V.R., and Ringach, D.L. (2007). Dynamics of receptive field size in primary visual cortex. J. Neurophysiol. *97*, 407–414.
42. Mazer, J.A., Vinje, W.E., McDermott, J., Schiller, P.H., and Gallant, J.L. (2002). Spatial frequency and orientation tuning dynamics in area V1. Proc. Natl. Acad. Sci. USA *99*, 1645–1650.
43. Casagrande, V.A., and Ichida, J.M. (2002). The lateral geniculate nucleus. In Physiology of the Eye, F.H. Adler, ed. (C.V. Mosby Company), pp. 655–668.
44. Taylor, M.M., and Creelman, C.D. (1967). PEST: efficient estimates on probability functions. J. Acoust. Soc. Am. *41*, 782–787.
45. Santini, F., Redner, G., Iovin, R., and Rucci, M. (2007). EyeRIS: a general-purpose system for eye-movement-contingent display control. Behav. Res. Methods *39*, 350–364.
46. Aytekin, M., Victor, J.D., and Rucci, M. (2014). The visual input to the retina during natural head-free fixation. J. Neurosci. *34*, 12701–12715.

Current Biology, Volume *27*

Supplemental Information

Consequences of the Oculomotor Cycle for the Dynamics of Perception

Marco Boi, Martina Poletti, Jonathan D. Victor, and Michele Rucci

Supplemental Experimental Procedures

Subjects. A total of 21 subjects with normal uncorrected vision took part in the study (age range 20-29; 17 Caucasian, 3 Asian, and one African American; 9 males and 12 females). Only emmetropic observers were enrolled in this study to ensure accurate gaze-contingent control of retinal stimulation. Fourteen observers participated in the recordings with natural images (Figures 1 and 2) and seven in the forced-choice experiments: four subjects in both Experiments 1 and 3 (Figures 4 and 6) and two in Experiment 2 (Figure 5). With the exception of one of the authors, all subjects were naïve about the purposes of the experiments and were paid to participate. Informed consent was obtained from all participants following the procedures approved by the Boston University Charles River Campus Institutional Review Board.

Apparatus and Stimuli. Stimuli were displayed on a gamma-corrected fast-phosphor CRT monitor (Iyama HM204DT) in a dimly-illuminated room. They were observed monocularly with the left eye patched, while movements of the right eye were recorded by means of a Dual Purkinje Image eyetracker (Fourward Technology) and sampled at 1 Khz. A dental imprint bite bar and a head-rest prevented head movements. Stimuli were rendered by means of EyeRIS, a custom system that enables precise synchronization between oculomotor events and the refresh of the image on the monitor [S1].

The data in Figures 1 and 2 were acquired as subjects freely examined grayscale pictures of natural scenes. Images were displayed sequentially, each for 10 s, at 100 Hz refresh rate and spatial resolution of 1024×768, so that the visual angle subtended by each pixel was similar to that of a pixel in the camera when the image was originally acquired ($\sim 1'$). Subjects were instructed to memorize the images.

In the forced-choice experiments (Experiments 1-3, Figures 4-6), stimuli consisted of Gabor patches with gratings tilted by $\pm 45^o$ ($\sigma = 61'$; spatial frequency either 1 or 10 cycles/degree). They were embedded within a full-screen naturalistic noise field in Experiment 1 (spectral density: k^{-2} in the range 0-24 cycles/deg; k spatial frequency) and, to specifically test the impact of eye movements at individual spatial frequencies, presented over a uniform background in Experiments 2 and 3. As expected, removal of the noise field increased sensitivity to both low and high spatial frequencies [S2] (cf. Figures 4 and 5). Both the average luminance of the noise field and the background luminance were 7 cd/m^2. Stimuli were displayed at a resolution of 800×600 pixel (Nyquist frequency: 23 cycles/deg) and vertical refresh rate of 200 Hz . In Experiments 2 and 3, fine changes in gray levels were obtained by means of a combination of bit stealing [S3] and dynamic dithering [S4]. There was no need for these techniques in Experiment 1, as the decrement in sensitivity caused by the background noise allowed reliable thresholds estimation

with the standard 8-bit gray-scale.

Procedure. Data were collected in many experimental sessions, each lasting approximately one hour. Every session started with preliminary setup operations that included positioning the subject in the apparatus, tuning the eyetracker, and calibrating EyeRIS to accurately convert the eye position measurements given by the eyetracker into screen coordinates. Subjects were never constrained in the experimental setup for blocks longer than 15 minutes.

In Experiments 1-3, subjects reported the orientation ($\pm 45^o$) of a grating in a forced-choice paradigm. Each trial started with the appearance of a fixation marker (a $12'$ red dot; Figure 4*A*) and, in Experiment 1, the noise field. After a random delay (1-1.5 s), the fixation marker shifted horizontally by $400'$, instructing the subject to saccade to this new location. The grating was then displayed centered at the cued location, embedded in the noise field when present. The grating's orientation and the noise pattern changed randomly at every trial. This randomization, along with use of fixed marker and cue positions, ensured similar attentional modulations across trials.

To expose the fovea to the natural luminance modulations caused by eye movements, visual stimulation was tightly coupled with oculomotor events. The stimulus was displayed on the first CRT frame (*i.e.*, within 10 ms) following real-time detection of instantaneous eye speed larger than 10^o/s and maintained for either 100 ms or 800 ms following the time at which the instantaneous eye speed dropped below 3^o/s. Subjects reported the grating's orientation by pressing different buttons on a joypad and received auditory feedback about their performance. All trials in which the saccade landed more than 1^o away from the center of stimulus or in which subjects did not give a response within 4 s were discarded. The contrast of the grating was adaptively adjusted at each trial following the PEST algorithm [S5].

Retinal stabilization in Experiment 2 was achieved by means of EyeRIS (see supplementary information in [S6] for a detailed description of stabilization accuracy). The stimulus was moved on the screen under real-time control to compensate for the subject's eye movements during the 800 ms period of post-saccadic exposure. Normal (unstabilized) and stabilized trials were randomly interleaved. Subjects did not report any stimulus fading under retinal stabilization and were in general unable to tell whether or not a trial was stabilized. Elimination of saccade transients in Experiment 3 was achieved by gradually increasing the grating's contrast from zero to the selected trial value for a period of 1.5 s starting at fixation onset. Following the contrast ramp, the grating was displayed at a fixed contrast for additional 800 ms.

Data analysis. Recorded traces were segmented into periods of fixation and saccades based on eye velocity. All movements with speeds larger than 3°/s and amplitudes exceeding $3'$ were classified as saccades. Saccade amplitude was defined as the length of the vector connecting the two locations at which velocity became greater (onset) and lower (offset) than 2^o/s. The segments in between successive saccades were labeled as periods of ocular drift. Only trials with accurate, uninterrupted tracking, in which the first and fourth Purkinje images always remained fully visible, were considered. To isolate temporal influences from the initial saccade and from ocular drifts, all trials with blinks, microsaccades, and/or uninstructed saccades were discarded from data analysis. In keeping with previous studies [S7, S8], microsaccades (saccades smaller than $30'$) were rare during free-viewing of natural scenes (Figure 1; average rate: 0.17 ± 0.15 microsaccade/s).

Frequency analyses relied on our recent spectral model of the retinal input [S9]. With standard nonparametric methods of power spectrum estimation, high frequency resolution can only be achieved using long windows of observation, an approach that conflicts with the naturally brief durations of saccades and fixation periods. To circumvent this problem, we used the displacement probability $q(\mathbf{x}, t)$, the probability that the eye moved by $\mathbf{x}$ in an interval t. Its Fourier Transform, $Q(\mathbf{k}, \omega)$ ($\mathbf{k}$ spatial frequency; ω temporal frequency) describes how, on average, the considered type of oculomotor activity redistributes spatial power in the joint space-time domain. Under the assumption that eye movements are statistically independent from the pattern of luminance—a very plausible assumption when considered across the entire visual field— multiplication of Q by the power spectrum of the external stimulus, $I(\mathbf{k})$, enables estimation of the frequency content of the retinal input:

$$S(\mathbf{k}, \omega) = I(\mathbf{k})\, Q(\mathbf{k}, \omega) \tag{1}$$

We have already shown empirically that this method gives excellent approximations of the power spectra of visual input signals estimated, at lower resolutions, by more standard approaches [S9].

The power redistribution function Q resulting from fixational drift was estimated over a total of 1,128 fixation periods, each longer than 512 ms (on average, 88 fixations per subject). The average power redistribution caused by saccades was estimated over 3,810 saccades with 1-10^o amplitude (on average, 272 saccades per subject). Saccade modulations were isolated by extracting each event from its original trace and placing it at the center of an artificial 512 ms trace in which the eye remained immobile before and after the saccade. In Figure 2, the power redistribution functions Q were multiplied by the ideal power spectrum of natural images [S10]—*i.e.*, scaled by $I(k) = k^{-2}$ in Eq. 1— to show how eye

movements transform the spatial power of natural scenes on the retina. To summarize results in two dimensions (space and time) spectra were radially averaged across spatial frequencies ($k = ||\mathbf{k}||$). All spectral analyses were carried out separately for each subject and then averaged across subjects.

Thresholds in Figures 4-6 were selected as the Michelson contrasts yielding 75% correct performance. They were estimated by fitting, via a maximum-likelihood procedure [S11], the data collected over at least four experimental sessions by means of a cumulative log-normal function [S12]. Only two subjects were tested in Figure 5 because of the massive amount of data needed from each subject to ensure high accuracy in retinal stabilization and the high consistency of results. For each subject, statistical significance in sensitivity differences was assessed by means of parametric bootstrap [S13] over N=2000 bootstrap trials. Specifically we tested whether the change in sensitivity $\Delta S = \log\left(\frac{S_1}{S_2}\right)$ differed from zero in the direction predicted by the theory, where S_1 and S_2 represent the sensitivities measured in the two conditions. Since our theory makes specific predictions about how experimental manipulations affect sensitivity, bootstrap tests are one-tailed. For the reader's convenience, in Figures 4 and 6, we also summarize our findings by reporting the results of paired t-tests applied to the log sensitivities across subjects. This use of t-tests is justified by previous reports that log sensitivity data conform to normal distributions [S14]. All reported t-tests are two-tailed.

Neural modeling. The mean instantaneous firing rates of simple cells in the primary visual cortex were modeled by means of standard spatiotemporal filters designed on the basis of neurophysiological data (Figure 1C), as in our previous models [S15–S17]:

$$r(t) = [I(\mathbf{x}, t) * K(\mathbf{x}, t) + N_I]_0 \tag{2}$$

where $[.]_0$ indicates rectification with zero threshold ($[x]_0 = x$ if $x > 0$, 0 otherwise) of the convolution between the retinal input $I(\mathbf{x}, t)$ and the cell kernel K, and N_I represents a Gaussian, zero-mean noise term. Filters were separable in their spatial and temporal components: $K(\mathbf{x}, t) = F(\mathbf{x})\ H(t)$. Spatial receptive fields were modeled as Gabor functions [S18] with peak frequency sensitivity of either 1 or 10 cycles/deg. The temporal response was a biphasic gamma function [S19]. Spatial receptive fields at different spatial frequencies were normalized to possess equal energy. Model neurons were stimulated by reconstructions of the visual input, $I(\mathbf{x}, t)$, experienced by the subjects in our experiments. In Figure 1F, receptive fields moved following recorded traces of eye movements over natural images. Data represents average responses over 500 fixations aligned by saccade end.

In the contrast detection model (Figure 3), the stimulus was a natural noise field $N_E(\mathbf{x})$ with or without a grating $G(\mathbf{x})$. The grating was always oriented at 45^o, possessed variable contrast C, and, in different experiments, frequency of either 1 or 10 cycles/deg. Neurons were tuned to the grating's frequency and orientation, and their receptive fields moved over the stimulus following traces composed of two fixations separated by a saccade. To model the dynamics of contrast sensitivity, a decision-making stage continually integrated neuronal responses starting from saccade end. In each trial, the grating was reported to be visible as soon as the integrated response, averaged across all simulated neurons ($\int_0^t \overline{r}(t)dt$) exceeded a threshold. At each time t during post-saccadic fixation, contrast thresholds were selected as the contrast values yielding hits and false alarm rates of 0.75 and 0.25, respectively ($d' = 1.35$). Thresholds in Figure 3E were estimated over $N = 500$ saccade-fixation sequences.

Supplemental References

S1. Santini, F., Redner, G., Iovin, R. & Rucci, M. (2007). EyeRIS: A general-purpose system for eye movement contingent display control. *Behav. Res. Methods* **39**, 350–364.

S2. Webster, M. A. & Miyahara, E. (1997). Contrast adaptation and the spatial structure of natural images. *J. Opt. Soc. Am.* **14**, 1–19.

S3. Tyler, C. W. (1997). Colour bit-stealing to enhance the luminance resolution of digital displays on a single pixel basis. *Spatial Vision* **10**, 369–77.

S4. Allard, R. & Faubert, J. (2008). The noisy-bit method for digital displays: converting a 256 luminance resolution into a continuous resolution. *Behav. Res. Methods* **40**, 735–743.

S5. Taylor, M. M. & Creelman, C. D. (1967). PEST: Efficient estimates on probability functions. *J. Acoust. Soc. Am.* **41**, 782–787.

S6. Rucci, M., Iovin, R., Poletti, M. & Santini, F. (2007). Miniature eye movements enhance fine spatial detail. *Nature* **447**, 852–855.

S7. Collewijn, H. & Kowler, E. (2008). The significance of microsaccades for vision and oculomotor control. *J. Vis.* **8**, 1–21.

S8. Poletti, M. & Rucci, M. (2016). A compact field guide to the study of microsaccades: Challenges and functions. *Vision Res.* **118**, 83–97.

S9. Kuang, X., Poletti, M., Victor, J. D. & Rucci, M. (2012). Temporal encoding of spatial information during active visual fixation. *Curr. Biol.* **22**, 510–514.

S10. Field, D. (1987). Relations between the statistics of natural images and the response properties of cortical cells. *J. Opt. Soc. Am. A* **4**, 2379–2394.

S11. Wichmann, F. A. & Hill, N. J. (2001). The psychometric function: I. fitting, sampling and goodness of fit. *Percept. Psychophys.* **63**, 1293–1313.

S12. Hall, J. L. (1981). Hybrid adaptive procedure for estimation of psychometric functions. *J. Acoust. Soc. Am.* **69**, 1763–1769.

S13. Wichmann, F. A. & Hill, N. J. (2001). The psychometric function: II. bootstrap-based confidence intervals and sampling. *Percept. Psychophys.* **63**, 1314–1329.

S14. Graham, N. V. S. (1989). *Visual pattern analyzers* (Oxford University Press, New York, 1989).

S15. Rucci, M., Edelman, G. & Wray, J. (2000). Modeling LGN responses during free-viewing: A possible role of microscopic eye movements in the refinement of cortical orientation selectivity. *J. Neurosci.* **20**, 4708–4720.

S16. Rucci, M. & Casile, A. (2005). Fixational instability and natural image statistics: Implications for early visual representations. *Network: Comp. Neural* **16**, 121–138.

S17. Desbordes, G. & Rucci, M. (2007). A model of the dynamics of retinal activity during natural visual fixation. *Visual Neurosci.* **24**, 217–230.

S18. Jones, J. P. & Palmer, L. A. (1987). An evaluation of the two-dimensional gabor filter model of simple receptive fields in cat striate cortex. *J. Neurophysiol.* **58**, 1233–1257.

S19. DeAngelis, G., Ohzawa, I. & Freeman, R. (1993). Spatiotemporal organization of simple-cell receptive fields in the cat's striate cortex. I. General characteristics and postnatal development. *J. Neurophysiol.* **69**, 1091–1117.