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.