

# Active Vision: Dynamic Reformatting of Visual Information by the Saccade-Drift Cycle

Igor Kagan<sup>1,\*</sup> and David C. Burr<sup>2</sup>

<sup>1</sup>Decision and Awareness Group, Cognitive Neuroscience Laboratory, German Primate Centre, Leibniz Institute for Primate Research, Goettingen 37077, Germany

<sup>2</sup>Department of Neuroscience, University of Florence, Italy

\*Correspondence: [ikagan@dpz.eu](mailto:ikagan@dpz.eu)

<http://dx.doi.org/10.1016/j.cub.2017.03.042>

Visual processing depends on rapid parsing of global features followed by analysis of fine detail. A new study suggests that this transformation is enabled by a cycle of saccades and fixational drifts, which reformat visual input to match the spatiotemporal sensitivity of fast and slow neuronal pathways.

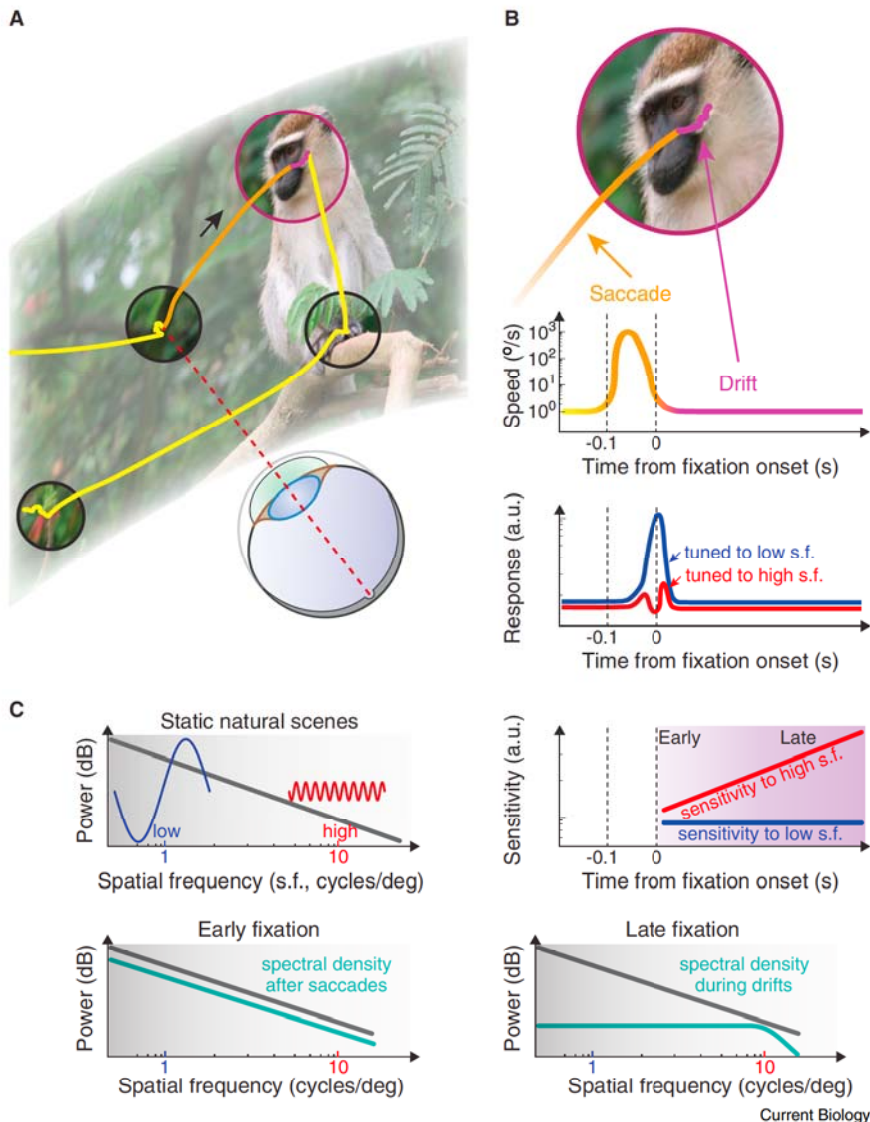
The first neuronal recordings in awake behaving monkeys performing eye movements were conducted more than 50 years ago, but most of what is known about the neural encoding in the early visual system in primates comes from experiments in anaesthetized or steadily fixating monkeys. However, it has been extensively documented that eye movements exert a strong influence on perception in humans [1]. Recently, the work of Rucci, Victor and colleagues [2,3] has demonstrated that smaller and slower eye movements occurring during fixation also affect visual processing, facilitating the perception of fine spatial features. Now a new study from the Rucci group reported in this issue of *Current Biology* [4] shows that both saccades and

slower fixational drifts reshape the spectral content of natural visual input to match the spatiotemporal profiles of fast magnocellular and slow parvocellular pathways. This dynamic reformatting could underlie a system of coarse-to-fine analysis of visual scenes, which has been proposed as an efficient means of image coding, often ascribed to the dynamic tuning properties of thalamic and cortical neurons in the magnocellular and parvocellular channels [5,6].

The notion that vision performs a progressive coarse-to-fine scene analysis — analysing progressively low to high spatial frequencies — goes back to the work of David Marr, and has been frequently revived, with empirical support [7,8]. The study of Boi *et al.* [4] gives a

new perspective on these theories, providing a plausible mechanism whereby normal oculomotor behaviour could underlie the cause of the temporal ordering of the analysis by fast magnocellular and slow parvocellular neurons, which respond preferentially to post-saccadic transients or in a sustained manner during drifts [9].

In contrast to previous work on coarse-to-fine perception of abruptly presented images during fixation, Boi *et al.* [4] recorded the eye movements of subjects freely exploring natural scenes in unconstrained conditions, and showed how these should, in theory, affect visual processing. They reported, as many others have noted, that the eyes are never still, but follow the well-known pattern of fast, ballistic



**Figure 1. Saccades and slow fixational drifts enable coarse-to-fine analysis of visual scenes.** (A) Eye position trace during fast saccades and slower drifts in periods of fixation (circles) during free viewing. (B) A one saccade (orange)–fixational drift (magenta) sequence. Top plot: retinal image speed as function of time; bottom plot: simulated firing rate of neurons tuned to low (blue curve) and high (red curve) spatial frequency in response to spatiotemporal luminance modulations imposed by eye movements. (C) Top left: ideal  $1/f^2$  ( $f$  – spatial frequency) power spectrum of static natural scenes, dominated by low frequencies. Top right: dynamics of post-saccadic contrast sensitivity to low (blue curve) and high (red curve) spatial frequency gratings embedded in noise. This pattern of contrast sensitivity results from a domination of the spectral power by low spatial frequencies immediately after saccades, across all non-zero temporal frequencies (early fixation, bottom left), and a subsequent redistribution of power across a broad range of spatial frequencies by ensuing drifts, which counterbalance the  $1/f^2$  spectrum of natural scenes, enhancing high spatial frequencies (late fixation, bottom right).

saccades at very fast speeds, interspersed with slow, seemingly random drifts of about  $1^\circ/\text{s}$ . Using the same time-varying inputs as experienced by free-viewing subjects, they simulated the response of two realistic classes of cortical simple cells, one tuned to low spatial frequencies

(1 cycle/ $^\circ$ ), the other to high spatial frequencies (10 cycles/ $^\circ$ ), both temporally selective to frequencies around 10 Hz. Given the temporal tuning, neither population responded well to stationary stimuli; however, both saccades and slow drifts generate a temporal modulation of the images,

which increases the simulated responses (Figure 1). The saccades modulate the image rapidly, and this favours the response to low spatial frequencies: as the temporal frequency of modulation is the product of speed and spatial frequency, the temporal modulation caused by saccades is preferred by cells with lower spatial tuning.

Thus, saccades cause a stronger response in the population selective for low spatial frequencies, with a peak towards the end of the saccade, at the onset of the fixation period. During the period of drift, however, the simulated responses of both cell classes were lower but similar, even though the presented images had 100 times more power in the low than in the high spatial frequencies, following the typical natural scenes spectral power distribution. This spatial frequency content is redistributed by temporal modulations due to the two types of eye movements in a distinct manner: saccades preserve the prevalence of low spatial frequencies, while drifts counterbalance this prevalence, thereby enhancing high spatial frequencies. The pattern of preferred sensitivities — to low spatial frequencies immediately after the saccade followed by a relative boost of high spatial frequencies — is consistent with the notion of coarse-to-fine analysis. Indeed, the simulations of post-saccadic time course of contrast sensitivity during ‘detection’ of noisy stimuli predicted that the detection of low spatial frequency is facilitated only immediately after the saccade, while that of high spatial frequency improves gradually during fixation.

Boi *et al.* [4] next tested the idea psychophysically, and found empirical validation for the major predictions: similar to the model, sensitivity to low spatial frequencies did not improve throughout fixation, while that to high spatial frequencies did, resulting in a relative shift in the weighting of sensitivity to the two frequencies. It has long been known that high frequency stationary stimuli benefit more from temporal summation than do low frequencies [10], but the present study provides the mechanism for the enhanced summation. Indeed, when Boi *et al.* [4] eliminated the slow drifts by retinal image

stabilization, the benefit of prolonged summation during fixation was far less. Conversely, the removal of post-saccadic transients decreased sensitivity to low but not to high spatial frequencies, demonstrating the importance of saccades for the perception of coarse features.

Most research into vision at the time of eye movements regards saccades as a difficult problem for visual system to solve [1]. But Rucci's approach has been to ask how vision can benefit from these movements. He and his team have previously shown how slow drifts can benefit sensitivity to high spatial frequencies [2,3,11], and now suggest that the saccades themselves can also be advantageous, by structuring the temporal ordering of analysis, so coarse information is analysed first, providing a 'scaffolding' for subsequent filling in with finer details.

Perhaps the most exciting suggestion of this new paper [4] is that eye movements — both fast and slow — serve to actively *reformat* the visual image, converting spatial information to temporal modulations on the retina. This concept of encoding space through time has been around for a long time [12], and re-emerges periodically [13–16]. Several demonstrations show that when objects are in motion, temporal signals can be used to provide information about space [13,15]: in effect, motion converts time to space. Some forty years ago, Horace Barlow wrote a provocative opinion piece suggesting that spatiotemporal interpolation could be fundamental for increasing spatial resolution past the photoreceptor sampling limit (see discussion in [13]). Rucci's group takes this idea a step further in showing that the image motion does not need to be caused by an external object — instead the eyes themselves may generate the motion necessary to convert space to time. And having two types of motion allows this to occur selectively for coarse and fine scales, implementing a hierarchical scheme of visual analysis. These results also suggest that distinct spatiotemporal characteristics of magnocellular and parvocellular neurons and the oculomotor behaviour evolved together to match the spectral content of the visual flow to the retina during natural viewing.

The study by Boi *et al.* [4] raises several intriguing questions. The first obvious question is how this study can be reconciled with evidence showing that saccades cause strong suppression of visual signals, particularly of low spatial frequencies ostensibly selective for the magnocellular system [1,17]. Although seemingly contradictory, there are several important differences between the two studies. Firstly, saccadic suppression is maximal at the beginning of the saccade, and an enhancement has often been reported after the saccade is concluded [1,9,17,18], in agreement with Boi *et al.*'s [4] findings showing maximal visual enhancement at fixation onset. Secondly, maximum suppression was observed at very low spatial frequencies, lower than 0.1 cycle/°, with little suppression at the moderate frequency of 1 cycle/° [1]. Thirdly, most work on saccadic suppression has been done with very large saccades (over 20°, whereas this study used small 7° saccades). And finally, there is evidence suggesting that it is specifically the motion signals that are suppressed during saccades [17], rather than the spatial information *per se*.

It is reasonable to suppose that saccades generate many processes, including both an enhancement of mid-range spatial frequencies and a suppression of motion driven by the very low spatial frequencies. Nevertheless, to distinguish between the physical, visual consequences of eye movements and centrally driven suppression due to extraretinal/efference copy signals [9], Rucci and colleagues should conduct similar measurements with 'simulated saccades', where the image is moved artificially at comparable speeds and durations [18]. Any differences observed between the real and simulated saccades could be put down to active suppression mechanisms. Similarly, it would be interesting to relate Boi *et al.*'s [4] novel take on the role of eye movements to the more classic work showing active processes accompanying saccades, such as compression and remapping [1].

Finally, it is important to ask to what extent is the reformatting of visual input goal-directed, and how adaptive is oculomotor behaviour for the purposes of task requirements or image statistics

of a scene? Is any of the process under voluntary control? It has been shown, also by Rucci and colleagues, that even the smallest fixational microsaccades can be accurately directed to the most sensitive part of the fovea to view the minute detail of interest [19]. Can slow drifts be also controlled precisely and adapted to the current requirements? There is considerable evidence for control of drifts [20], but the neural mechanisms for such control are still unknown.

## REFERENCES

- 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.
- Rucci, M., Iovin, R., Poletti, M., and Santini, F. (2007). Miniature eye movements enhance fine spatial detail. *Nature* 447, 852–855.
- 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.
- Boi, M., Poletti, M., Victor, J.D., and Rucci, M. (2017). Consequences of the oculomotor cycle for the dynamics of perception. *Curr. Biol.* 27, 1268–1277.
- 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.
- Kauffmann, L., Ramanoël, S., and Peyrin, C. (2014). The neural bases of spatial frequency processing during scene perception. *Front. Integr. Neurosci.* 8.
- Watt, R.J. (1987). Scanning from coarse to fine spatial scales in the human visual system after the onset of a stimulus. *JOSA A* 4, 2006–2021.
- 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.
- 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, 1–19.
- Burr, D.C. (1981). Temporal summation of moving images by the human visual system. *Proc. R. Soc. Lond. B* 211, 321–339.
- Kagan, I. (2012). Active vision: fixational eye movements help seeing space in Time. *Curr. Biol.* 22, R186–R188.
- Marshall, W.H., and Talbot, S.A. (1942). Recent evidence for neural mechanisms in vision leading to a general theory of sensory acuity. In *Visual Mechanisms* (Oxford, England: Jacques Cattell), pp. 117–164.
- Burr, D., and Ross, J. (1986). Visual processing of motion. *Trends Neurosci.* 9, 304–307.

14. Ahissar, E., and Arieli, A. (2001). Figuring space by time. *Neuron* 32, 185–201.
  15. Nishida, S. (2004). Motion-based analysis of spatial patterns by the human visual system. *Curr. Biol.* 14, 830–839.
  16. Burr, D., and Ross, J. (2004). Vision: the World through picket fences. *Curr. Biol.* 14, R381–R382.
  17. Burr, D.C., Holt, J., Johnstone, J.R., and Ross, J. (1982). Selective depression of motion sensitivity during saccades. *J. Physiol.* 333, 1–15.
  18. Diamond, M.R., Ross, J., and Morrone, M.C. (2000). Extraretinal control of saccadic suppression. *J. Neurosci.* 20, 3449–3455.
  19. Poletti, M., Listorti, C., and Rucci, M. (2013). Microscopic eye movements compensate for nonhomogeneous vision within the fovea. *Curr. Biol.* 23, 1691–1695.
  20. Rucci, M., and Poletti, M. (2015). Control and functions of fixational eye movements. *Annu. Rev. Vis. Sci.* 1, 499–518.
-