



Published in final edited form as:

Trends Cogn Sci. 2018 October ; 22(10): 883–895. doi:10.1016/j.tics.2018.07.009.

Temporal Coding of Visual Space

Michele Rucci^{1,2}, Ehud Ahissar³, and David Burr^{4,5,*}

¹Department of Center for Visual Science, University of Rochester, Rochester, NY 14627

²Department of Brain and Cognitive Sciences, University of Rochester, Rochester, NY 14627.

³Department of Neurobiology, Weizmann Institute, Rehovot, Israel

⁴Department of Neuroscience, University of Florence, Florence 50125, Italy

⁵School of Psychology, University of Sydney, Camperdown, 2006, Australia

Abstract

Establishing a representation of space is a major goal for sensory systems. Spatial information, however, is not always explicit in the incoming sensory signals. In most modalities, it needs to be actively extracted from cues embedded in the temporal flow of receptor activation. Vision, on the other hand, starts with a sophisticated optical imaging system that explicitly preserves spatial information on the retina. This may lead to the assumption that vision is predominantly a spatial process: all that is needed is to transmit the retinal image to the cortex, like uploading a digital photograph, to establish a spatial map of the world. However, this deceptively simple analogy is inconsistent with theoretical models and experiments that study visual processing in the context of normal motor behavior. We argue here that like other senses, vision relies heavily on temporal strategies and temporal neural codes to extract and represent spatial information.

Keywords

eye movements; space perception; temporal processing; visual system; retina; visual fixation; saccade; microsaccade; ocular drift

Stable visual representations, a moving visual world

Like a camera, the eye forms an image of the external scene on its posterior surface where the retina is located, with its dense mosaic of **photoreceptors** that convert light into electrochemical signals. At each moment in time, all spatial information is present in the visual signals striking the photoreceptors, which explicitly encode space by their position within the retinal array. This **camera model of the eye** and the **spatial coding** idea have long dominated visual neuroscience. Although the specific reference frames of spatial

*Corresponding author.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

representations (e.g., retinotopic vs. spatiotopic) have been intensely debated [1,2], spatial information has always been assumed to originate from the receptor layout in the retina.

Alas, the eye does not behave like a camera. While a photographer usually takes great care to ensure that the camera does not move, the eyes insist on moving continuously [3–6]. Humans perform rapid gaze shifts, known as **saccades**, two - three times per second. And even though models of the visual system often assume the visual input to be a stationary image during **fixational pauses** between successive saccades, small eye movements, known as **fixational eye movements**, continually occur. These movements displace the stimulus by considerable amounts on the retina, therefore continually changing the light signals striking the photoreceptors ([7,8]; see Box 1).

Furthermore, unlike the film in a camera, the visual system depends on temporal transients. Neurons in the retina, thalamus, and later stages of the visual pathways respond strongly to temporal changes [9–13]. Visual percepts tend to fade away in the complete absence of temporal transients [14–17], and spatial changes that occur too slowly are not even detected by humans [18,19].

These considerations do not appear compatible with the standard idea that space is encoded solely by the position of neurons within spatial maps. They suggest that the visual system combines spatial sampling with temporal processing to extract and encode spatial information. In this article, we describe specific mechanisms by which this process may take place and briefly review supporting evidence. We specifically focus on the input signals resulting from eye movements, as these induce the most common temporal changes experienced by the retina.

The need for spatio-temporal codes

Most agree that at some level the visual system must integrate spatial information over time. Unlike the uniform resolution of a photograph acquired by a conventional camera, the image projected on the retina is sampled by the neural array in a highly non-homogeneous manner. Only in a very small region at the center of gaze, the **fovea**, are cones packed tightly enough to allow for the high resolution we enjoy. Acuity drops rapidly as images fall outside this region [20,21]; and even within the fovea, cone density [22] and visual function [23] are not uniform, and attention exerts its influence [24]. To perceive the fine spatial structure of the world, we therefore need to center gaze on the area of interest, by making saccadic eye movements (see Box 1).

Saccades, of all sizes, are critical for visual perception: they are as essential for vision as is the focusing of the image on the retina. But they also present serious challenges to purely spatial models of vision, as each saccade causes abrupt changes to the retinal images, and to their centrally projected signals. Not only do these frequent movements shift the image around - like a video filmed by a careening drunk - they also change the cortical distortion, as the center of cortical magnification travels across different parts of the visual scene. The successive snapshot of each new fixation would be uninterpretable without taking eye position into account, and high-level spatial representations require temporal integration

across saccades. How the visual system achieves these goals has been subject to intense debate over the last few decades, well-reviewed in many recent publications [25,26].

In this essay, we concentrate not so much on the problems generated by eye- movements, but on their benefits and how they could contribute to schemes of **temporal coding** and **spatiotemporal coding**. It is well known that the spatial acuity of human vision can vastly exceed that predicted by the retinal sampling density, even in the fovea. For example, vernier acuity (the smallest perceivable offset between two lines or dots) - often termed hyperacuity - can be as small as a few arcseconds, an order of magnitude smaller than photoreceptor spacing [27]. While it is commonly assumed that this level of resolution may be achieved by spatial summation mechanisms in the visual cortex, simple calculations under the actual conditions of human vision show that pure spatial integration could never lead to the actual observed resolution (Box 2). Thus, we are compelled to consider other aspects of the image, such as its temporal structure.

It is remarkable that the visual system can establish clear representations of fine spatial patterns at all, given how much the eyes actually move during the acquisition of visual information [6–8,28]. As described in Box 1, eye movements occur incessantly even when trying hard to maintain fixation on an object. In the fixational pauses, the eyes wander erratically around with a motion known as **ocular drift**, occasionally superimposed with miniature replicas of saccades, termed **microsaccades** [6–8] (see Box 1). Although the amplitudes of the ocular drift fluctuations are close to the **eye-tracking resolution** limits of most popular devices, they are large enough to displace the stimulus over tens of foveal receptors, producing motion signals that would be immediately perceived if caused by the motion of objects in the scene, rather than eye movements. Since neurons in the retina, thalamus, and visual cortices often integrate signals over temporal windows of 100–150 ms and more [12,13,29,30], fundamental questions emerge about why the world does not appear blurred and how the visual system can reliably represent fine spatial detail [31–34].

So how is fine spatial vision possible? An intriguing hypothesis, raised several times over the course of almost a century [35–43], is that, rather than being detrimental, eye movements may actually be beneficial for encoding space, particularly at small scales. This happens because spatial information is not lost in the visual signals resulting from eye movements, but converted into modulations that may be extracted via temporal codes.

Although this concept may at first appear counter-intuitive, one clear example of how temporal information can aid spatial analysis is motion perception (Box 3). Objects of interest are often in motion, presenting a problem for purely spatial mechanisms of neural encoding, as the windows of neuronal integration are long enough for even moderate speeds of motion to smear images beyond recognition [30]. Yet this does not happen: images in motion do not appear to be particularly smeared, provided they are displayed for a long enough periods of time [32]. Clearly the visual system is specialized to analyze moving objects, taking their movements into account while extracting information about them [42,44] (more details in Box 3).

Motion perception is often considered as a specialized visual function that relies on dedicated neural machinery, separate from that normally used to process stationary spatial scenes. But although objects may be stationary in space, eye movements keep their projections in continuous motion on the retina. We argue here that motion processing is not a special case for vision, but the norm. As there are no stationary retinal signals during natural vision, motion processing is the fundamental, basic operating mode of human vision. Thus, the visual system must always process space via temporal signals, especially those resulting from eye movements.

Eye movements, like object motion, link space and time

A well-established finding, preserved across species and types of neurons, is that retinal, thalamic, and early visual cortical neurons respond much more strongly to changing than stationary stimuli [12,13,45–47]. In the laboratory, these neurons are commonly activated by temporally modulating stimuli on the display (either by flashing or moving them), but under natural viewing conditions, the most common cause of temporal modulation on the retina is our own behavior: moving our eyes. Eye movements continually transform stationary spatial scenes into a spatiotemporal **luminance flow** on the retina. When in the appropriate frequency bands, these spatio-temporal patterns can potentially strongly activate neurons along the visual system.

A helpful way to conceptualize and analyze the space-time reformatting resulting from eye movements is by means of the spectral distribution of the luminance flow, a representation of the power of the retinal stimulus over **spatial frequencies** and **temporal frequencies** (Fig. 1A). When a static scene is observed with immobile eyes — a situation that never actually occurs under natural conditions — the input to the retina is a static image. Its power is confined to 0 Hz (green line in Fig. 1A). Eye movements transform this static scene into a spatiotemporal flow, an operation that, in the frequency domain, is equivalent to redistributing the 0 Hz power across nonzero temporal frequencies (blue surface in Fig. 1A). The exact way in which this redistribution occurs depends on the specific characteristics of the eye movements: because of their very different properties, saccades and drifts create signals with different spatiotemporal statistics on the retina.

Effect of eye drift on visual input signals

Let us first examine the spectral redistribution (the change in the input **power spectrum**) caused by ocular drift. Drift-induced retinal modulations depend on the spatial frequency of the stimulus. This happens because, as the eye moves, the luminance signals impinging onto the retina tend to fluctuate rapidly with stimuli at high spatial frequency — stimuli that contain sharp edges and texture — but will remain more constant with smooth, low spatial frequency stimuli.

This intuition is quantitatively captured by the blue curve in Fig. 1B, which shows the proportion of spatial power in the otherwise static image (the 0 Hz power) that drifts shift to non-zero temporal frequencies. For simplicity of visualization, Fig. 1B shows a section of the input power at 10 Hz (averaged across all spatial directions), but sections at different nonzero temporal frequencies will be similar. The amount of power that shifts to nonzero

temporal frequencies increases proportionally with the square of the spatial frequency, up to a limit that varies with the considered temporal frequency (approximately 15 cycles/deg for 10 Hz). Power tends to grow less rapidly at higher spatial frequencies, but when summing up all temporal sections within the range of human sensitivity, an enhancement can be observed up to 30 cycles/deg, close to the spatial resolution of the photoreceptor array [37].

The space-time reformatting caused by ocular drift combines with the spectral density of natural images in an interesting way [48]. The spectral power in natural scenes declines approximately with the square of the spatial frequency [49], resulting in less power at high spatial frequencies, the opposite to the amplification caused by ocular drift. Thus, the motion of the eye enhances the frequency components that have lower power in natural scenes. These two effects counterbalance each other, so that fixational modulations at different spatial frequencies contain approximately equal power up to the range at which the amplification by fixational eye movements begins to be attenuated (Fig. 1C).

This equalization - known as spectral **whitening** - reveals a form of matching between normal drift and the natural world statistics. Whitening would happen for any Brownian-like eye drift pattern, but there is a relationship between the speed of the drift (*i.e.*, its effective diffusion constant) and the spatial frequency up to which whitening occurs. In humans, whitening occurs only up to the spatial frequencies that can be sampled by the photoreceptors, indicating that drift motion is well matched to the characteristics of the retina.

In sum, ocular drift causes a very specific input reformatting during normal fixation, transforming natural scenes with a spectral density that is heavily biased to low spatial frequencies (the $1/k^2$ line in Fig. 1C) into luminance fluctuations with equalized spatial power at nonzero temporal frequencies. This is the dominant input signal during the late phase of inter-saccadic fixation, when neurons are no longer affected by the consequences of the preceding saccade.

Importantly, spectral whitening of the visual flow is not just a laboratory phenomenon that only occurs when the head is restrained - a common condition in experiments in which small eye movements are recorded. When the head is free to move normally, fixational head instability also occurs, but drifts partly compensate for this motion [6,50]. As a consequence, the residual retinal image motion delivered by the interaction between fixational head and eye movements maintains spectral distributions almost identical to those measured with the head constrained [50,51].

Effect of saccades and microsaccades on visual input signals

Under natural conditions, ocular drift does not occur in isolation, but alternates with rapid gaze shifts of various magnitudes, from microsaccades (within-fovea shifts), to saccades (Box 1). These movements, both large and small, position objects of interest on the small high-acuity region of the retina known as the fovea [23,52,53]. But in doing so they also produce sharp temporal transients to the retina, which likely drive neuronal responses in the early phase of fixation, immediately following each saccade.

As one may expect, the power redistribution caused by these rapid eye movements differs considerably from that caused by ocular drift. Unlike fixational drift modulations, these abrupt changes redistribute temporal power in a similar way across a broad range of spatial frequencies (red curve in Fig. 1B), preserving the spectral distributions of the external scene. This implies that during viewing of natural scenes the power of the visual input to the retina will decline with the square of spatial frequency, like the spectrum of natural images, at all relevant non-zero temporal frequencies (see distributions at 10 Hz in Fig. 1C). Because of this effect, saccades transiently generate more temporally-modulated power than drift at low spatial frequencies, and comparable power around 10 cycles/deg.

In sum, during natural viewing, the normal alternation between saccades and fixational drifts sequentially exposes neurons to modulations with different kinds of spatial information. At fixation onset, neurons respond to an input signal dominated by low spatial frequencies: later during fixation, the ocular drift enhances power at high spatial frequencies, and the same neurons are now driven by a visual flow with equal power across spatial frequencies (Fig. 1C).

Perceptual consequences of oculomotor-induced temporal modulations

The arguments above clearly show that the standard assumption that the input to the retina is a stationary image is a gross simplification. In reality, neurons are exposed to a complex spatiotemporal flow that cyclically changes statistics during normal eye movements. But do these input modulations have real and important consequences to perception?

Consider first the visual signals originating from ocular drift. These modulations are predicted to enhance retinal responses to high spatial frequencies. This hypothesis can be tested experimentally by a technique known as **retinal stabilization**, which attempts to eliminate - and in practice drastically reduces - retinal image motion, by counter-moving the stimulus in real time against the eye movement [5,54,55]. Consistent with predictions, recent evidence shows that eliminating drift movements selectively impairs sensitivity at high spatial frequencies [56–58].

Figure 1D compares contrast sensitivity measured for an observer under normal viewing, and with drift movements stabilized. As expected from the space-time reformatting shown in Fig. 1D, reduction of drift movements strongly decreases sensitivity at 10 cycles/deg, but has virtually no impact at 1 cycle/deg. Interestingly, these findings conflict with the widespread belief that fixational drift is primarily helpful to prevent image fading at low spatial frequencies [15,59,60]. On the contrary, human vision takes advantage of drift modulations to enhance detailed spatial information, providing strong support to long standing dynamic theories of visual acuity [36].

The power redistributions in Fig. 1C suggest that saccades boost low spatial frequencies relative to fixational drift. This prediction is confirmed by experiments in which the normal visual transient caused by a saccade is eliminated. The data in Fig. 1E summarize results from an experiment in which subjects performed a large saccade under two conditions: in the first, the stimulus (a grating at either 1 or 10 cycles/deg) appeared as soon as the saccade

was detected, to expose the fovea to the normal transients elicited by the saccade; in the other the stimulus was slowly faded in after the end of the saccade, to completely eliminate the saccade-generated transient. Elimination of the saccade transient strongly reduces sensitivity at 1 cycle/deg, but has virtually no effect at 10 cycles/deg. Thus, consistent with the power redistributions of Fig. 1B, at low spatial frequencies, modulations from saccades and drifts cannot be traded, as the latter are not beneficial in this range, regardless of their duration.

A coarse-to-fine strategy of visual analysis

These considerations carry important implications regarding the formation of spatial representations, and their dynamics. Because retinal neurons are cyclically exposed to temporal transients with different spatial content, one may expect the perceptual process responsible for establishing spatial representations to be tuned to the oculomotor cycle. To optimally accrue spatial information from the oculomotor-shaped retinal input flow, an observer should rely heavily on saccade transients to extract low spatial frequencies, and integrate this information throughout the course of fixation with spatial frequency information. In other words, low spatial frequencies should be processed before high spatial frequencies.

The visual system appears to follow these dynamics. As shown in Fig. 1F, sensitivity to low spatial frequency quickly reaches its plateau shortly after the end of a saccade and changes little during the course of fixation; in contrast, sensitivity to high spatial frequencies can increase more than twofold over a one-second interval [58]. This leads to a sequential coarse-to-fine analysis during each post-saccadic fixation, first the coarser structure given by low spatial frequencies, followed by a fine-scale analysis of higher spatial frequencies. This strategy may enable tuning of visual processing via control of eye movements, for example by performing shorter fixational pauses in low spatial frequency tasks, and could help the system integrate successive fixations, locking the low-frequency information that can be resolved outside central vision.

That mechanisms tuned to high spatial frequencies have longer processing times than those tuned to low has long been known. It goes back to early ideas of “sustained and transient” channels [30,61]. This different timing also maps into the anatomically based distinction of magnocellular and parvocellular streams (or M- and P-streams), with the M-stream preferring low spatial and high-temporal frequencies, and the P-stream vice versa [47]. But why should neurons sensitive to separate spatial frequency ranges also differ in their temporal characteristics?

Our observations suggest an answer to this question. In a system designed to encode space through time, analysis of low but not high spatial frequencies tends to be more reliable immediately following a saccade than later during fixation. Distinct temporal strategies are, therefore, necessary for efficient acquisition of spatial information in different frequency ranges, and the spatiotemporal characteristics of magnocellular and parvocellular neurons are well suited to carry out these functions. In keeping with this proposal, it has been shown in macaque V1 that neurons that are activated by drifts have smaller receptive fields, longer

response latency and prefer slower stimulus motion compared with neurons transiently activated by microsaccades and saccades, with larger receptive fields and faster responses [46].

Possible strategies for space-time encoding

The previous considerations highlight the importance of eye movements in transforming a static spatial scene into a spatiotemporal flow on the retina and some of the perceptual consequences. But what are the neural mechanisms that take advantage of this transformation? Neurophysiological investigations have only recently started to examine the beneficial consequences of eye movements [46,62–73], but multiple theoretical possibilities emerge. Given how closely the temporal power in the retinal flow predicts visual responses, and the close relationship between power and temporal correlation (the power spectrum is the Fourier Transform of the autocorrelation function), an obvious candidate is the structure of correlated neural activity in the early stages of the visual system, particularly zero-lag correlations, since synchronous neuronal firing is already known to critically enhance the efficacy of neural propagation [74,75].

The spectral distributions in Fig. 1C directly predict that the saccade-drift alternation would (a) first elicit broad pools of simultaneously active neurons immediately after a saccade; and (b) that these pools progressively shrink during fixation to eventually yield a pattern of uncorrelated activity (a direct consequence of the “whitened” flow resulting from eye drift) during viewing of natural scenes [76]. Since the synchronous neuronal responses would signal the contrast changes induced by eye movements, they would first encode the general low-spatial frequency structure of the visual scene immediately after a saccade; and they would then convey the pattern of edges (local changes in contrast) later during fixation [77]. These same mechanisms could be responsible for representing the gist of the scene immediately following the first few saccades.

Equalizing the power of a signal across spatial frequencies is equivalent to discarding spatial correlations, reducing the intrinsic redundancy of natural scenes, which many theories of vision regard as essential for efficient encoding [78–80]. What is interesting, however, is that the normal drift patterns of the eyes, ensures that this decorrelation occurs in the input signals themselves, before any neural processing takes place. And indeed, physiological recordings in vitro support the proposal that fixational instability reduces the extent of correlated activity relative to abrupt onset of the stimulus [68]. This operation removes the average pairwise correlations present in natural scenes, while leaving higher-order correlations, such as edges. It discards components of the image predictable from the general structure of natural scenes, thus emphasizing the features of the specific image under examination.

The use of synchronous modulations is only one of the many strategies by which the visual system can extract, encode, and decode spatial information from the incoming spatio-temporal flow. Encoding temporal strategies can take multiple forms. One basic form is the scheme of frequency coding. Since eye movements convert spatial offsets into temporal delays, two contrast changes separated in the image by a distance dx will activate a given

photoreceptor (or ganglion cell) sequentially, with a delay of dt ($dt=dx/v$, where v is eye velocity). Thus, spatial frequencies can be signaled via the temporal frequencies of spikes [40,41,81]. Another form of temporal encoding is the scheme of phase coding: two receptors (or ganglion cells) will be activated with a delay dt proportional to the spatial offset dx between the contrast changes they face.

These schemes allow high-resolution coding of shape, texture and motion [40]. Importantly, in these schemes, spatial resolution would depend on the resolution of neuronal temporal processing, rather than on the spatial granularity of the receptors layout on the retina, providing a possible explanation for the mechanisms of fine spatial vision and hyperacuity [40,41]. Furthermore, as hypothesized by some of us, the visual system could take advantage of the temporal power generated by eye movements to synchronize cortical circuits (oscillators) engaged in thalamocortical phase-locked loops [40,82–84]. Ocular drift provides strong modulations in the low-temporal frequency range [40,85,86]; locking to these modulations would be beneficial, during fixational pauses, for decoding the incoming flow and extracting spatial patterns and motion information [40].

Concluding remarks and future perspectives

In this brief essay we suggest an alternative, non-conventional view of how the visual system encodes and represents space. Rather than relying exclusively on the standard spatial coding scheme in which spatial information is conveyed solely by the position of neurons within maps, we argue that spatial representations primarily rely on the temporal signals resulting from self-motion, particularly eye movements. Furthermore, we argue that this strategy of converting space to time likely employs mechanisms that are also responsible for the perception of visual motion.

Multiple predictions emerge, some of which are summarized in Outstanding Questions. Depending on the specific neural mechanisms responsible for encoding temporal signals, there may exist distinct coding patterns for shape, texture and motion. For example, fine detail of shape may be encoded primarily by inter-receptor temporal phase (representing external spatial phase), texture primarily by instantaneous intra-burst frequencies of individual receptors (representing external spatial frequencies), and motion by inter-burst temporal frequencies (representing Doppler shifts generated by the relative motion of the eye and external objects) [40].

Furthermore, while in this essay we have specifically focused on the consequences of eye movements on the retinal input, the mechanisms of spatial encoding may benefit from extraretinal information for fixational eye movements, as both microsaccades [23,52,87] and ocular drift [50,51,88,89] are under control. If, on the contrary, extraretinal copies of ocular drift are not available or not used, one would expect (a) artificially generated temporal lags to be perceived as spatial offsets; and (b) that jittered full-field scenes (unlike non-coherent motion of parts of the scene) should go unnoticed without impairing spatial vision and hyperacuity.

The more general conclusion from the considerations discussed here is that primates plan their various eye movements not just to position objects of interest on the fovea, but also to generate the temporal transients that most effectively emphasize the spatial scales relevant to the task. These and other emerging hypotheses will be the focus of future investigations.

Acknowledgements

Supported by NIH R01 EY018363 and NSF grants BCS-1457283, the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 786949) and ECSPLAIN-FP7- IDEAS-ERC-ref. 338866. We thank Alessandro Benedetto for technical assistance.

Glossary

Camera model of the eye

A model in which the eye operates like a camera, taking discrete spatial snapshots of the scene, frame after frame.

Eye-tracking resolution

Fixational eye movements reach fluctuation amplitudes that are beyond the reach of many eye-tracking devices. Analyses of eye movements should take the noise level and sensitivity of the measuring device into account to determine the lower bounds in the estimated kinematic variables.

Fixational eye movements

small eye movements during fixational pauses. They are traditionally subdivided in microsaccades, ocular drift and tremor. However, the distinction between drift and tremor is challenging and not attempted in this article.

Fixational pause

the period in between saccades, in which the stimulus moves relatively little on the retina. Often termed fixation for brevity

Fovea

here we use this term to refer to the foveola, the small (~0.2 mm in diameter, approximately 1 degree in visual angle) retinal region at the center of gaze free from blood vessels and rod photoreceptors where cones are most densely packed

Luminance flow on the retina

the continually changing pattern of light impinging onto the retina even during viewing of a stationary scene.

Microsaccade

a term that has been used to describe both the small saccades that, during normal viewing, maintain the stimulus of interest within the fovea, and the involuntary small saccades that occur in the laboratory when observers are instructed to maintain steady gaze on a marker. Microsaccades properties are continuous with those of saccades and do not form a distinct group. Furthermore, recent evidence indicates that microsaccades act as small saccades, initiating a new fixational pause.

Minute of arc (or arcmin, or minarc, or')

a measurement unit of angle, corresponding to one-sixtieth of one degree. It is usually indicated by the symbol

Ocular drift

the incessant wandering of the eye during fixational pauses. Its power spectrum declines with the square of temporal frequency, often exhibiting a mode around the gamma (~30–80 Hz) frequency band. This latter component is occasionally termed “tremor”

Photoreceptors

The neurons in the retina that convert light into electrochemical signals

Power spectrum

a representation of how the power of a signal is distributed in the frequency domain

Retinal ganglion cells

the neurons in the output stage of the retina, which relay spike-based information to brainstem, midbrain and thalamic areas.

Retinal stabilization

a laboratory procedure that aims to eliminate the physiological motion of the retinal image

Saccade

The most rapid eye movement normally used to bring visual regions of interest to the foveola. Saccades typically occur 2–3 times per second

Spatial coding

a coding scheme in which the information is carried by the spatial location of the receptor and the intensity of its activation.

Spatial frequency

the number of times a sinusoidal pattern (a grating) repeats itself within one degree

Spatio-temporal coding

a coding scheme in which the information is carried by the relationships between the spatial locations of the receptors and the timing of their activations

Spatio-temporal receptive field

The region of space-time that can modify the response of a neural unit

Temporal coding

a coding scheme in which the information is carried by the timing of receptor activation

Temporal frequency

the number of times a sinusoidal modulation repeats itself within one second

Whitening

an operation that equalizes the power of a signal across all its frequencies (practically, within a given range). This process also removes any internal correlations in the signal

BIBLIOGRAPHY:

1. Burr DC and Morrone MC (2011) Spatiotopic coding and remapping in humans. *Philos. Trans. R. Soc. B Biol. Sci.* 366, 504–515
2. Hall NJ and Colby CL (2011) Remapping for visual stability. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 366, 528–39 [PubMed: 21242141]
3. Barlow HB (1952) Eye movements during fixation. *J. Physiol.* 116, 290–306 [PubMed: 14939180]
4. Ratliff F and Riggs LA (1950) Involuntary motions of the eye during monocular fixation. *J. Exp. Psychol.* 40, 687–701 [PubMed: 14803643]
5. Ditchburn RW (1973) *Eye-movements and visual perception*, Oxford Uni.
6. Steinman RM et al. (1973) Miniature Eye Movement. *Science* (80-.). 181, 810–819 [PubMed: 4198903]
7. Kowler E (2011) Eye movements: The past 25 years. *Vision Res.* 51,1457–1483 [PubMed: 21237189]
8. Rucci M and Poletti M (2015) Control and Functions of Fixational Eye Movements. *Annu. Rev. Vis. Sci.* 1, 499–518 [PubMed: 27795997]
9. Robson JG (1966) Spatial and Temporal Contrast-Sensitivity Functions of the Visual System. *J. Opt. Soc. Am.* 56,1141
10. Nagano T (1980) Temporal sensitivity of the human visual system to sinusoidal gratings. *J. Opt. Soc. Am.* 70, 711–6 [PubMed: 7400873]
11. 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 [PubMed: 1136158]
12. Benardete EA and Kaplan E (1999) The dynamics of primate M retinal ganglion cells. *Vis. Neurosci.* 16, 355–368 [PubMed: 10367969]
13. Benardete EA and Kaplan E (1997) The receptive field of the primate P retinal ganglion cell, II: Nonlinear dynamics. *Vis. Neurosci.* 14,187–205 [PubMed: 9057279]
14. Troxler IPV (1804) über das Verschwinden gegebener Gegenstände innerhalb unseres Gesichtskreises. *Ophthalmol. Bibl.* 2,1–53
15. Kelly DH (1979) Motion and vision I Stabilized images of stationary gratings. *J. Opt. Soc. Am.* 69,1266 [PubMed: 521857]
16. Poletti M and Rucci M (2010) Eye movements under various conditions of image fading. *J. Vis.* 10,1–18
17. Steinman RM et al. (1985) Vision in the presence of known natural retinal image motion. *J. Opt. Soc. Am. A* 2, 226 [PubMed: 3973754]
18. Rensink RA (2002) Change Detection. *Annu. Rev. Psychol.* 53, 245–277 [PubMed: 11752486]
19. O'Regan JK et al. (1999) Change-blindness as a result of 'mudsplashes.' *Nature* 398, 34–34 [PubMed: 10078528]
20. Weymouth FW et al. (1928) Visual Acuity within the Area Centralis and its Relation to Eye Movements and Fixation. *Am.J. Ophthalmol.* 11, 947–960
21. Jacobs RJ (1979) Visual resolution and contour interaction in the fovea and periphery. *Vision Res.* 19,1187–1195 [PubMed: 550578]
22. Curcio CA et al. (1990) Human photoreceptor topography. *J. Comp. Neurol.* 292, 497–523 [PubMed: 2324310]
23. Poletti M et al. (2013) Microscopic Eye Movements Compensate for Nonhomogeneous Vision within the Fovea. *Curr. Biol.* 23,1691–1695 [PubMed: 23954428]
24. Poletti M et al. (2017) Selective attention within the foveola. *Nat. Neurosci.* 20,1413–1417 [PubMed: 28805816]
25. Ross J et al. (2001) Changes in visual perception at the time of saccades. *Trends Neurosci.* 24,113–121 [PubMed: 11164942]
26. Wurtz RH (2008) Neuronal mechanisms of visual stability. *Vision Res.* 48, 2070–2089 [PubMed: 18513781]

27. Westheimer G (2009) Hyperacuity In Encyclopedia of Neuroscience (Squire LR, ed), pp. 45–50, Academic Press
28. Cherici C et al. (2012) Precision of sustained fixation in trained and untrained observers. *J. Vis.* 12, 31–31 [PubMed: 22728680]
29. Barlow HB (1957) Increment thresholds at low intensities considered as signal/noise discriminations. *J. Physiol.* 136, 469–488
30. Burr DC (1981) Temporal summation of moving images by the human visual system. *Proc. R. Soc. London. Ser. B, Biol Sci.* 211, 321–39 [PubMed: 6111803]
31. Barlow HB (1979) Reconstructing the visual image in space and time. *Nature* 279,189–190 [PubMed: 440428]
32. Burr D (1980) Motion smear. *Nature* 284,164–165 [PubMed: 7360241]
33. Burak Y et al. (2010) Bayesian model of dynamic image stabilization in the visual system. *Proc. Natl. Acad. Sci. U. S. A.* 107,19525–30 [PubMed: 20937893]
34. Packer O and Williams DR (1992) Blurring by fixational eye movements. *Vision Res.* 32,1931–1939 [PubMed: 1287989]
35. Averill HL and Weymouth FW (1925) Visual perception and the retinal mosaic. II. The influence of eye-movements on the displacement threshold. *J. Comp. Psychol.* 5,147–176
36. Marshall WH and Talbot SA (1942) Recent evidence for neural mechanisms in vision leading to a general theory of sensory acuity. In *Biological Symposia—Visual Mechanisms* (Kluver, H., ed), pp. 117–164, Jacques Cattell
37. Rucci M and Victor JD (2015) The unsteady eye: an information processing stage, not a bug. *Trends Neurosci.* 38,195–206 [PubMed: 25698649]
38. Rucci M (2008) Fixational eye movements, natural image statistics, and fine spatial vision. *Netw. Comput. Neural Syst.* 19, 253–285
39. Rucci M and Casile A (2005) Fixational instability and natural image statistics: Implications for early visual representations. *Netw. Comput. Neural Syst.* 16,121–138
40. Ahissar E and Arieli A (2012) Seeing via Miniature Eye Movements: A Dynamic Hypothesis for Vision. *Front. Comput. Neurosci.* 6, 89 [PubMed: 23162458]
41. Ahissar E and Arieli A (2001) Figuring Space by Time. *Neuron* 32,185–201 [PubMed: 11683990]
42. Burr DC et al. (1986) Seeing objects in motion. *Proc. R. Soc. London. Ser. B, Biol. Sci.* 227, 249–65 [PubMed: 2871557]
43. Burr DC (1979) Acuity for apparent vernier offset. *Vision Res.* 19, 835–7 [PubMed: 483604]
44. Burr D and Ross J (1986) Visual processing of motion. *Trends Neurosci.* 9, 304–307
45. Croner LJ and Kaplan E (1995) Receptive fields of P and M ganglion cells across the primate retina. *Vision Res.* 35, 7–24 [PubMed: 7839612]
46. Kagan I et al. (2008) Saccades and drifts differentially modulate neuronal activity in V1: Effects of retinal image motion, position, and extraretinal influences. *J. Vis.* 8,19–19
47. Derrington AM and Lennie P (1984) Spatial and temporal contrast sensitivities of neurones in lateral geniculate nucleus of macaque. *J. Physiol.* 357, 219–40 [PubMed: 6512690]
48. Kuang X et al. (2012) Temporal Encoding of Spatial Information during Active Visual Fixation. *Curr. Biol* 22, 510–514 [PubMed: 22342751]
49. Field DJ (1987) Relations between the statistics of natural images and the response properties of cortical cells. *J. Opt. Soc. Am. A* 4, 2379 [PubMed: 3430225]
50. Poletti M et al. (2015) Head-Eye Coordination at a Microscopic Scale. *Curr. Biol.* 25, 3253–3259 [PubMed: 26687623]
51. Aytekin M et al. (2014) The visual input to the retina during natural head- free fixation. *J. Neurosci.* 34,12701–15 [PubMed: 25232108]
52. Ko H-K et al. (2010) Microsaccades precisely relocate gaze in a high visual acuity task. *Nat. Neurosci.* 13,1549–53 [PubMed: 21037583]
53. Poletti M and Rucci M (2016) A compact field guide to the study of microsaccades: Challenges and functions. *Vision Res.* 118, 83–97 [PubMed: 25689315]

54. Kelly DH (1981) Disappearance of stabilized chromatic gratings. *Science* 214,1257–8 [PubMed: 7302596]
55. Santini F et al. (2007) EyeRIS: A general-purpose system for eye- movement-contingent display control. *Behav. Res. Methods* 39, 350–364 [PubMed: 17958145]
56. Rucci M et al. (2007) Miniature eye movements enhance fine spatial detail. *Nature* 447, 852–855
57. Ratnam K et al. (2017) Benefits of retinal image motion at the limits of spatial vision. *J. Vis.* 17, 30 [PubMed: 28129414]
58. Boi M et al. (2017) Consequences of the Oculomotor Cycle for the Dynamics of Perception. *Curr. Biol* 27,1268–1277 [PubMed: 28434862]
59. Koenderink JJ (1972) Contrast Enhancement and the Negative Afterimage. *J. Opt. Soc. Am.* 62, 685 [PubMed: 5030998]
60. Tulunay-Keese U (1982) Fading of stabilized retinal images. *J. Opt. Soc. Am.* 72, 440–7 [PubMed: 7077431]
61. Kulikowski JJ and Tolhurst DJ (1973) Psychophysical evidence for sustained and transient detectors in human vision. *Physiol.* 232,149–62
62. Greschner M et al. (2002) Retinal ganglion cell synchronization by fixational eye movements improves feature estimation. *Nat. Neurosci.* 5, 341–347 [PubMed: 11914721]
63. Krauzlis RJ et al. (2017) Neuronal control of fixation and fixational eye movements. *Philos. Trans. R. Soc. B Biol. Sci.* 372, 20160205
64. Leopold DA and Logothetis NK (1998) Microsaccades differentially modulate neural activity in the striate and extrastriate visual cortex. *Exp. brain Res.* 123, 341–5 [PubMed: 9860273]
65. Herrington TM et al. (2009) The Effect of Microsaccades on the Correlation between Neural Activity and Behavior in Middle Temporal, Ventral Intraparietal, and Lateral Intraparietal Areas. *J. Neurosci.* 29, 5793–5805 [PubMed: 19420247]
66. Martinez-Conde S et al. (2004) The role of fixational eye movements in visual perception. *Nat. Rev. Neurosci.* 5, 229–240 [PubMed: 14976522]
67. Ibbotson M and Krekelberg B (2011) Visual perception and saccadic eye movements. *Curr. Opin. Neurobiol.* 21, 553–558 [PubMed: 21646014]
68. Segal IY et al. (2015) Decorrelation of retinal response to natural scenes by fixational eye movements. *Proc. Natl. Acad. Sci. U. S. A.* 112, 3110–5 [PubMed: 25713370]
69. Snodderly DM et al. (2001) Selective activation of visual cortex neurons by fixational eye movements: Implications for neural coding. *Vis. Neurosci.* 18, S0952523801182118
70. Gur M et al. (1997) Response variability of neurons in primary visual cortex (V1) of alert monkeys. *J. Neurosci.* 17, 2914–20 [PubMed: 9092612]
71. Snodderly DM (2016) A physiological perspective on fixational eye movements. *Vision Res.* 118, 31–47 [PubMed: 25536465]
72. Hafed ZM et al. (2009) A Neural Mechanism for Microsaccade Generation in the Primate Superior Colliculus. *Science* (80-.). 323, 940–943 [PubMed: 19213919]
73. Chen C-Y et al. (2015) Neuronal Response Gain Enhancement prior to Microsaccades. *Curr. Biol.* 25, 2065–2074 [PubMed: 26190072]
74. Dan Y et al. (1998) Coding of visual information by precisely correlated spikes in the lateral geniculate nucleus. *Nat. Neurosci.* 1, 501–507 [PubMed: 10196548]
75. Bruno RM and Sakmann B (2006) Cortex is driven by weak but synchronously active thalamocortical synapses. *Science* 312, 1622–7 [PubMed: 16778049]
76. Desbordes G and Rucci M (2007) A model of the dynamics of retinal activity during natural visual fixation. *Vis. Neurosci.* 24, 217–230 [PubMed: 17640413]
77. Rucci M and Casile A (2004) Decorrelation of neural activity during fixational instability: Possible implications for the refinement of V1 receptive fields. *Vis. Neurosci.* 21, 725–738 [PubMed: 15683560]
78. Attneave F (1954) Some informational aspects of visual perception. *Psychol. Rev.* 61,183–193 [PubMed: 13167245]
79. Barlow HB (2012) Possible Principles Underlying the Transformations of Sensory Messages. In *Sensory Communication* pp. 216–234, The MIT Press

80. Atick JJ and Redlich AN (1992) What Does the Retina Know about Natural Scenes? *Neural Comput.* 4,196–210
81. Ahissar E et al. (2016) On the possible roles of microsaccades and drifts in visual perception. *Vision Res.* 118, 25–30 [PubMed: 25535005]
82. Ahissar E (1998) Temporal-Code to Rate-Code Conversion by Neuronal Phase-Locked Loops. *Neural Comput.* 10, 597–650 [PubMed: 9527836]
83. Ahissar E and Vaadia E (1990) Oscillatory activity of single units in a somatosensory cortex of an awake monkey and their possible role in texture analysis. *Proc. Natl. Acad. Sci. U. S. A.* 87, 8935–9 [PubMed: 2247469]
84. Benedetto A and Morrone MC (2017) Saccadic Suppression Is Embedded Within Extended Oscillatory Modulation of Sensitivity. *J. Neurosci.* 37, 3661–3670 [PubMed: 28270573]
85. Bengi H and Thomas JG (1973) Studies on Human Ocular Tremor In *Perspectives in Biomedical Engineering* pp. 281–292, Palgrave Macmillan UK
86. Herrmann CJJ et al. (2017) A self-avoiding walk with neural delays as a model of fixational eye movements. *Sci. Rep.* 7,12958
87. Havermann K et al. (2014) Fine-scale plasticity of microscopic saccades. *J. Neurosci.* 34,11665–72 [PubMed: 25164662]
88. Nachmias J (1961) Determiners of the drift of the eye during monocular fixation. *Josa* 51, 761–766
89. Steinman RM (1995) Moveo ergo video: natural retinal image motion and its effect on vision In *Exploratory Vision: The Active Eye* (Landy MS *et al*, eds), pp. 3–50, Springer
90. Legge GE and Kersten D (1987) Contrast discrimination in peripheral vision. *J. Opt. Soc. Am. A.* 4,1594–8 [PubMed: 3625341]
91. Hansen T et al. (2009) Color perception in the intermediate periphery of the visual field. *J. Vis.* 9, 26–26
92. Nandy AS and Tjan BS (2012) Saccade-confounded image statistics explain visual crowding. *Nat. Neurosci.* 15, 463–469 [PubMed: 22231425]
93. Cornsweet TN (1956) Determination of the Stimuli for Involuntary Drifts and Saccadic Eye Movements*. *J. Opt. Soc. Am.* 46, 987 [PubMed: 13367941]
94. Fiorentini A and Ercoles AM (1966) Involuntary eye movements during attempted monocular fixation. *Atti Fond. Giorgio Ronchi* 21,199–217
95. Adler FH and Fliegelman M (1934) Influence of fixation on the visual acuity. *Arch. Ophthalmol* 12, 475–483
96. Ezenman M et al. (1985) Power spectra for ocular drift and tremor. *Vision Res.* 25,1635–1640 [PubMed: 3832587]
97. Ko H et al. (2016) Eye movements between saccades: Measuring ocular drift and tremor. *Vision Res.* 122, 93–104 [PubMed: 27068415]
98. Skavenski AA et al. (1979) Quality of retinal image stabilization during small natural and artificial body rotations in man. *Vision Res.* 19, 675–83 [PubMed: 547477]
99. Crick FHC et al. (1981) An Information Processing Approach to Understanding the Visual Cortex. In *Cortex The* (Schmidt EO, ed), MIT Press
100. Shapley R and Victor J (1986) Hyperacuity in cat retinal ganglion cells. *Science* 231, 999–1002 [PubMed: 3945816]
101. Reichardt W (1957) Autokorrelationsauswertung als Funktionsprinzip des Zentralnervensystems. *Zeitschriftfur Naturforsch.* 12, 447–457
102. Adelson EH and Bergen JR (1985) Spatiotemporal energy models for the perception of motion. *J. Opt. Soc. Am. A.* 2, 284–99 [PubMed: 3973762]
103. Watson AB and Ahumada AJ (1985) Model of human visual-motion sensing. *J. Opt. Soc. Am. A.* 2, 322–41 [PubMed: 3973764]
104. van Santen JPH and Sperling G (1985) Elaborated Reichardt detectors. *J. Opt. Soc. Am. A* 2, 300 [PubMed: 3973763]
105. Burr D and Thompson P (2011) Motion psychophysics: 1985–2010. *Vision Res.* 51,1431–1456 [PubMed: 21324335]

BOX 1.**Exploring via eye movements**

The human retina collects visual information from a large portion of space, a region that covers more than 180 degrees along the horizontal meridian. To effectively monitor such a large visual field, visual acuity and other functions are not uniform across the retina, but progressively deteriorate with increasing distance from the fovea [20,21,90–92], the region of the retina where cones are most densely packed. This tiny area, approximately the size of the full moon, occupies less than 0.01% of the visual field. Thus, to efficiently examine a visual scene, humans must move their eyes.

Figure I shows a typical pattern of eye movements performed by a human observer examining a stationary scene. The primary way that the fovea is reoriented toward objects of interest is via rapid gaze shifts, known as saccades, that normally occur 2–3 times every second. Depending on the task, saccades come in a broad range of amplitudes, ranging from a few **minutes of arc** to many degrees. Small saccades, often termed microsaccades, frequently occur during examination of fine spatial patterns, precisely centering the stimulus to optimize fine spatial vision [23,52].

In the intervals in between saccades and microsaccades, the eyes move incessantly with seemingly erratic trajectories. This motion is often considered as the superposition of two separate processes: a slower wandering, often referred to as ocular drift [93,94] and a faster jittery oscillation known as tremor [95–97]. In this article, we do not attempt to subdivide the intersaccadic motion of the eye, which we refer to as “ocular drift”, or simply “drift”. Ocular drift moves the eye in ways that resembles Brownian motion, frequently changing direction and maintaining the image on the retina in motion at speeds that would be immediately noticeable if resulting from the motion of objects in the scene rather than eye movements.

When the observer’s head is immobilized, a standard condition for measuring very small eye movements, the average speed of the stimulus on the retina is approximately 1°/s. The mean instantaneous speed increases (1.5°/s) when the head is free to move normally, but this increment is only a fraction of what one would expect, considering that the eyes now drift more than three times faster relative to when the head is immobilized. This happens because most of the increase in fixational eye movements acts to counterbalance fixational head movements [50,51,98], maintaining retinal image motion within levels similar to those experienced with the head constrained.

BOX 2.**Boundaries for the applicability of the camera model to vision**

A strong argument against purely spatial mechanisms of visual encoding comes from hyperacuity, the ability to detect relative offsets between lines or dots an order of magnitude smaller than the spatial granularity of the eye (~3" compared with 20"). A spatial model could account for such hyperacuity only by intensity-coding (sub-pixel interpolation). This could be done by optical or neural smearing, spreading visual detail across more than a single photoreceptor, allowing the intensities of retinal outputs to be interpolated to achieve super-resolution. Although there are enough cortical neurons for the required interpolation [31,99], the problem with any intensity model is that the build-up of an intensity code to allow such accurate interpolation takes time, and the eye is not steady during that time.

Intensity coding of a spatial offset (e.g. a vernier stimulus) is based on the difference in the number of spikes emitted by two populations of **retinal ganglion cells**. With finite firing rates, the difference in population spike-count [dS] induced by a spatial offset (dx , Fig. IA) increases monotonically with the duration over which spikes are counted (T). At first approximation, dS depends on dx and the maximal firing rate (R) but not on the population size (Appendix 3 in [40]):

$$dS = RT dp = 3RT dx$$

where dp is the spatial offset in units of individual photoreceptors at the center of the fovea (3 receptors per arcminute) and dx is in arcminutes.

The minimally discriminable difference is $dS = 1$ spike, the absolute lower bound for any intensity-based discrimination. Figure IB demonstrates the codependency of this bound on T and dx (red curve) for $R = 50$ spikes/s. The blue curve describes the mean expected diffusion distance of the gaze (D) due to ocular drift as a function of count duration (T), approximating ocular drift to a random walk process:

$$D = (VT)^{1/2}$$

where V is the speed of a single random-walk step, estimated here to be 0.16 '/ms.

For reliable intensity-based coding, gaze should drift less than dx during counting. Thus, the operating range is where $ds > 1$ and $dx > D$. Hence, with $R=50$ spikes/s, a firing difference of 1 spike can be generated only for spatial offsets $> 1'$, offsets at least an order of magnitude larger than vernier hyperacuity (~0.05'). To illustrate the implausibility of intensity-based coding of hyperacuity note that in order to obtain a 1-spike difference for a vernier stimulus (Fig. IA, black arrowheads) with such retinal motion the retinal maximal rate should be $R > 400,000$ spikes/s.

Thus, although this estimate varies with drift characteristics, pure intensity coding does not seem to be able to account for vernier acuity and probably not even for Snellen acuity. That mammalian visual systems do not use pure spatial coding is also supported

by the finding that hyperacuity is substantially poorer than what is predicted from pure spatial processing, when assuming stationary eyes [100]. These considerations point at **temporal coding** as the most feasible candidate for high-resolution retinal coding and hyperacuity.

BOX 3.**The problem of perceiving moving images without blur**

Any motion-detection system must sample over both space and time. The simplest motion detector often termed a *Reichardt detector*, samples at least two points in space, with samples delayed over time [101]. More in general, motion selectivity implies the presence of spatio-temporal receptive fields orientated in space-time, as illustrated in Fig. I. This concept of spatio-temporal filtering and spatio-temporal receptive fields forms the basis of many standard models of motion perception [42,102–105].

The slant of the receptive field clearly means that it will be selective to velocity, responding well only to trajectories oriented near its long axis. However, this very selectivity for motion makes the receptive field ideal to analyze the spatial form of moving objects. As the field is aligned to the motion trajectory, it essentially annuls to motion, as if it were tracking it. The analysis is now not only a spatial analysis, but a spatio-temporal analysis, orthogonal to the long-axis of the field's orientation. It will integrate signals within the neural unit as they traverse the field, without the blur that may be expected from a receptor array without temporal tuning.

If motion is not smooth, but stroboscopic (such as the cinema or television) it will be perceived as smooth. The system will interpolate between the motion samples. Clear evidence for spatio-temporal interpolation has been provided. If moving sampled images are perturbed only in time and not in space, by introducing a temporal offset in two vernier line segments, the system perceives the temporal offset as a spatial offset (see demo in on-line movie). And it does so with very similar acuity to real spatial offset. Put another way, a spatio-temporal receptive field, which traverses time and space will respond to changes in either dimension: that is to say, they allow motion to effectively convert time to space.

Highlights:

- The input to the visual system is not a static image: incessant eye movements continuously reformat input signals, transforming spatial patterns into temporal modulations on the retina.
- Luminance modulations resulting from different types of eye movements have different spatiotemporal frequency distributions. Saccade transients enhance lower spatial frequencies than those emphasized by modulations by fixational drift.
- Recent experiments that manipulated oculomotor transients indicate that humans are sensitive to these temporal modulations. The natural alternation between saccade and fixational drift sets the stage for a coarse-to-fine analysis of the scene.
- Encoding of spatial information in the visual system does not just rely on the position of neurons within maps, but also on temporal processing of the luminance modulations resulting from eye movements.

Outstanding questions

- How is visual information acquired? (I) When does acquisition occurs? (a) Upon landing of a saccade on a new region of interest, (b) during the fixational pause, or (c) during both? (II) How does acquisition occur? (a) Via spatial coding (the “camera model”, which we argue strongly against), (b) via temporal coding, or (c) via spatiotemporal coding?
- How is visual information processed? Addressing this question is entangled with major sub-questions: What makes the world perceived as stable in spite of sensor motion? Is processing based on open- or closed- loop environment-brain dynamics? Is central processing locked to eye movements, and if so to which components? Do spatial perception and motion perception rely on similar mechanisms?
- Do primates actively control eye movements to take advantage of their temporal transients? In principle, eye movements could be controlled independently from foveation needs to select the scale of visual processing, emphasize task-relevant ranges of spatial details and determine the resolution of temporal coding. It is presently unknown whether humans and other primates employ this form of active control and, if so, what are the controlled variables.
- Do saccades establish a clock for visual decoding? To effectively make use of the spatial information contained in the luminance modulations resulting from eye movements, analysis of neural activity needs to be timed relative to the occurrence of saccades. For example, a spike from the same neuron carries a different informational value if occurring during early fixation, when the input change caused by the preceding saccade still exerts its influence, or later, when temporal changes are only caused by ocular drift.

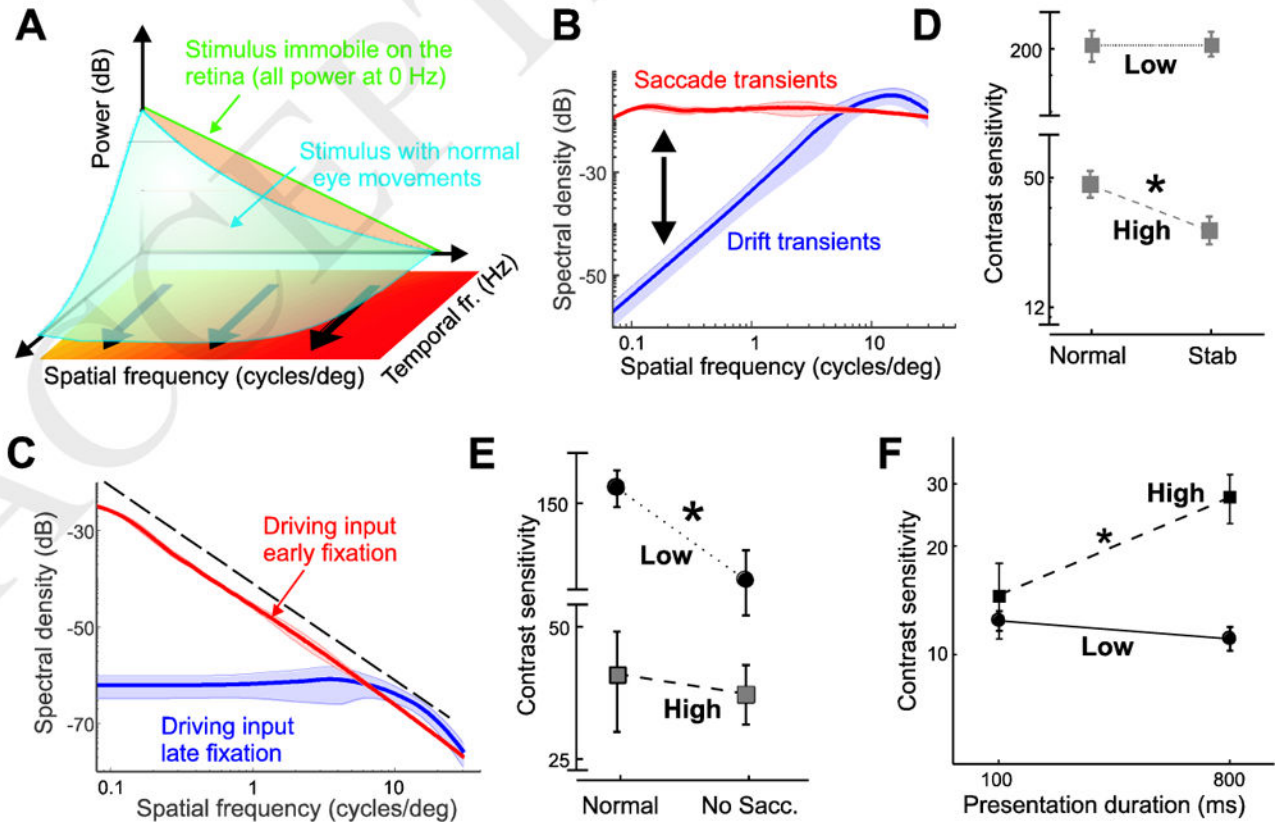


Figure 1. Space-time consequences of eye movements.

(A) Eye movements transform spatial patterns of luminance into temporal modulations impinging on the retina. In the frequency domain, this transformation redistributes the spatial power of the stimulus from 0 Hz (green line) across nonzero temporal frequencies (blue surface) in a way that depends on the type of eye movements. (B) Spatiotemporal power redistributions resulting from saccades (red curve) and fixational drifts (blue). Each curve represents the average fraction of stimulus power that eye movements make available at 10 Hz. (C) Spectral density of the retinal input during viewing of natural scenes. Data represent average distributions at 10 Hz. The power distributions made available by saccades and drift drive cell responses at different times during the course of post-saccadic fixation. (D-E) Drift and saccade transients contribute to different ranges of visual sensitivity. Elimination of drift and saccade transients (D) impairs contrast sensitivity at 10 cycles/deg (high), but has little effect at 1 cycle/deg (low). The opposite effect occurs with elimination of saccade transients (E). (F) Time course of visual sensitivity during natural post-saccadic fixation. Following a saccade, contrast sensitivity continues to improve at 10 cycles/deg but not at 1 cycle/deg. Data adapted from [58]. (*) indicates $p < 0.03$.

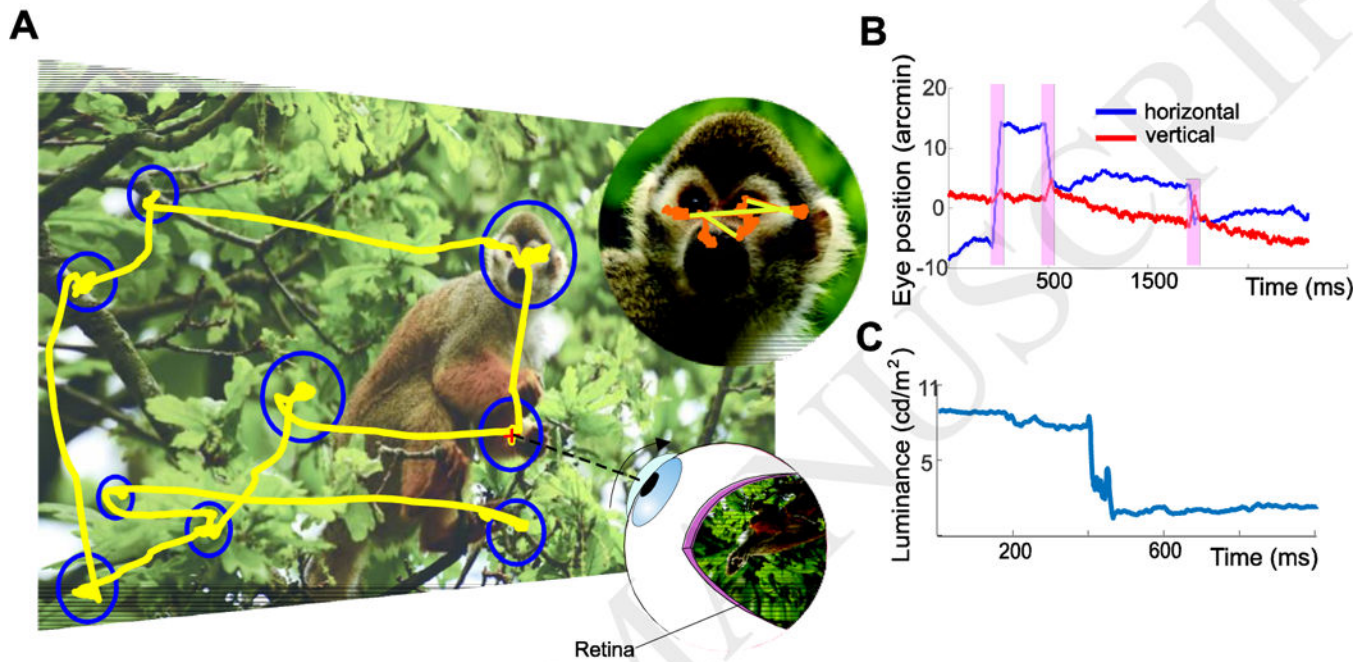


Figure I (Box 1). Normal eye movements.

(A) An oculomotor trace (yellow) superimposed on the observed scene. Each fixational pause is represented by a circle with radius proportional to its duration. The zoomed-in panel (top right) shows fixational eye movements. Both microsaccades (yellow segments) and drifts (orange) are visible. (B) Horizontal and vertical eye displacements in a portion of the trace. The shaded red regions indicate microsaccades. (C) Luminance modulations experienced by one retinal receptor, as the eye changes fixation via a saccade.

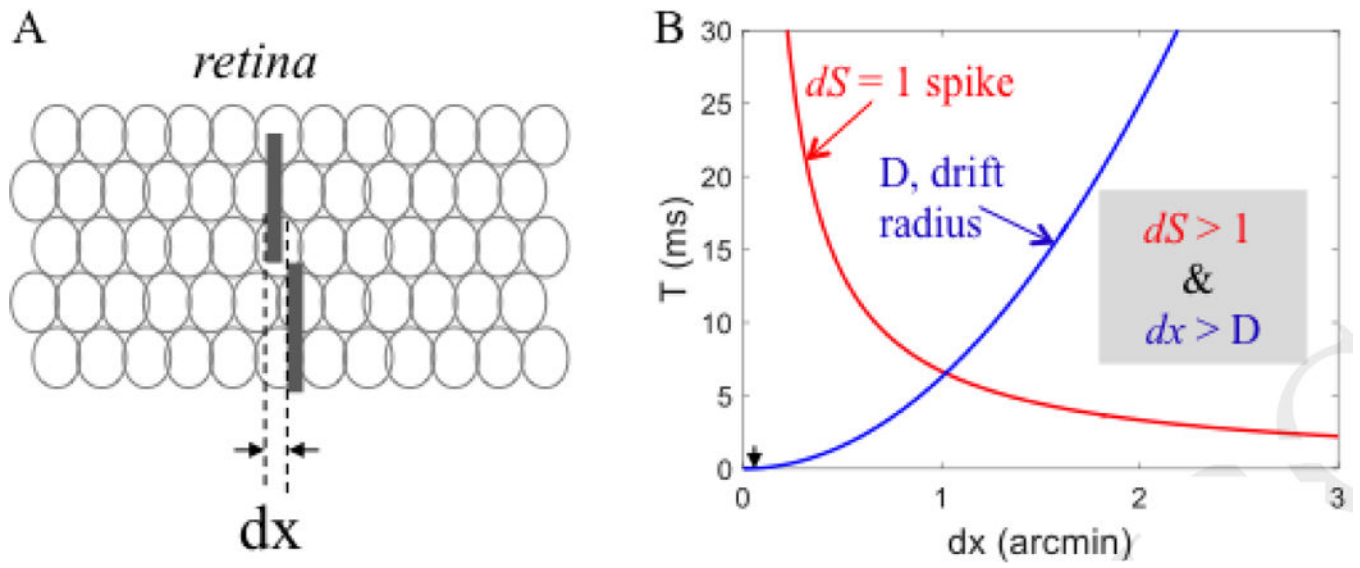


Figure I (Box 2).
Boundaries for spatial coding of vernier offsets.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

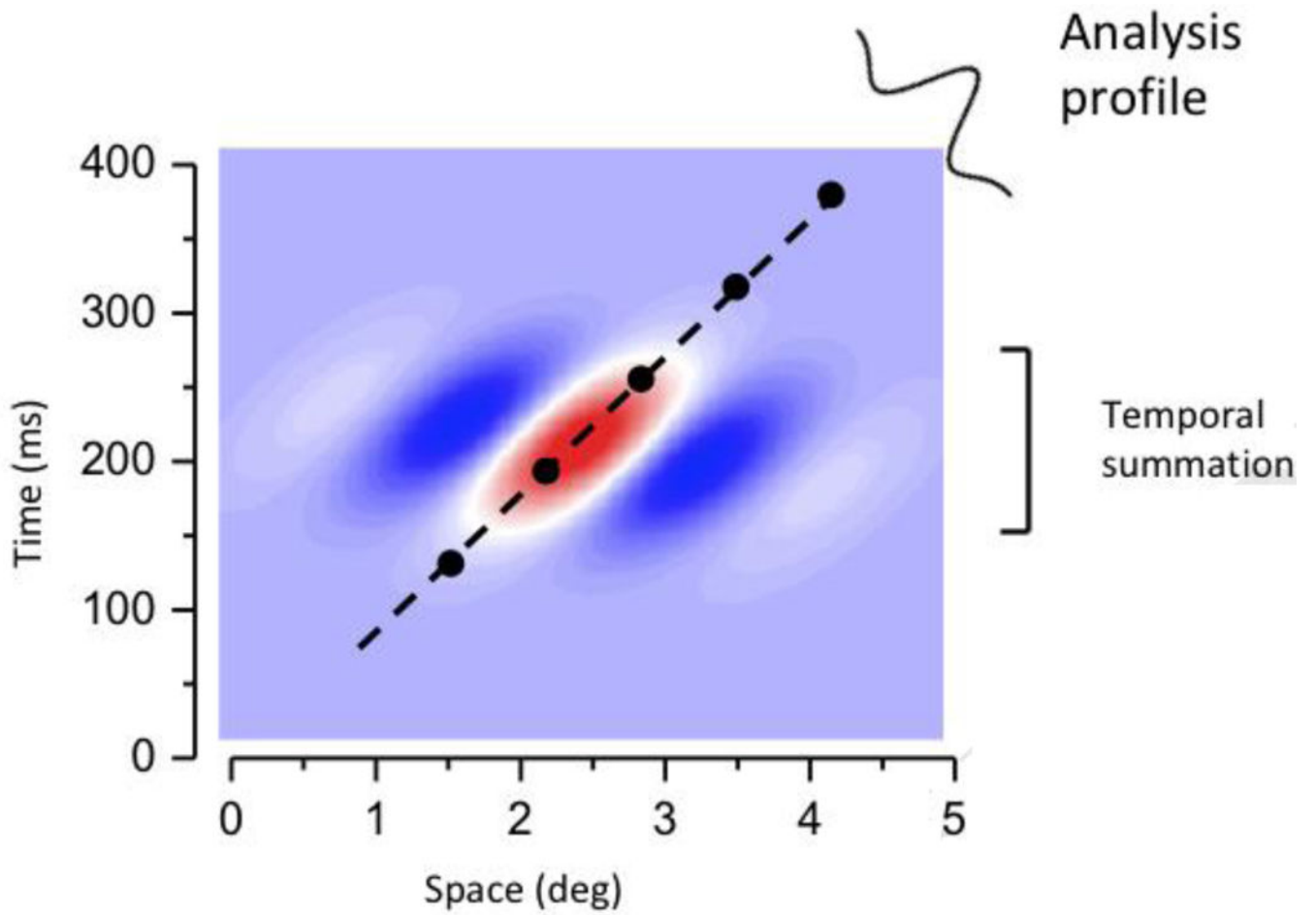


Figure I (Box 3): Spatio-temporal sensitivity profile of a unit selective to rightward motion. Red indicates positive response, blue negative.