Article

PERCEPTION

Perception
2015, Vol. 44(8–9) 986–994

DOI: 10.1177/0301006615594946
pec.sagepub.com
SAGE

1-D Vision: Encoding of Eye Movements by Simple Receptive Fields

Ehud Ahissar, Shira Ozana and Amos Arieli
Department of Neurobiology, Weizmann Institute of Science,
Rehovot, Israel

Abstract
Eye movements (eyeM) are an essential component of visual perception. They allow the sampling and scanning of stationary scenes at various spatial scales, primarily at the scene level, via saccades, and at the local level, via fixational eyeM. Given the constant motion of visual images on the retina, a crucial factor in resolving spatial ambiguities related to the external scene is the exact trajectory of eyeM. We show here that the trajectory of eyeM can be encoded at high resolution by simple retinal receptive fields of the symmetrical type. We also show that such encoding can account for motion illusions such as the Ouchi illusion. In addition, encoding of motion projections *along* horizontal and vertical symmetrical simple retinal receptive fields entails a kind of Cartesian decomposition of the 2-D image into two 1-D projections.

Keywords
active vision, fixational eye movements, drift, simple cells, perception, visual stability

During natural vision, the eyes are always on the move. These eye movements (eyeM) are the "engine" enabling vision, at least of stationary scenes. If the eyes would not move, the change-seeking photoreceptors would hardly be activated by a stationary scene. However, such dynamic vision comes with a computational cost, if not for the visual system than for us, scientists who try to understand it. The cost is in the complexity of retinal coding; the retina cannot be considered an analog of a stationary camera in which the output contains a spatial copy of the external image. Instead, the retinal output contains dynamic patterns of activity whose spatiotemporal relationships have to be decoded in relation to the eyeM that generated them (Ahissar & Arieli, 2001, 2012; Burak, Rokni, Meister, & Sompolinsky, 2010). This requires a neuronal registration of the on-going eyeM. Note that without the registration of eyeM, ambiguity is expected with any scheme of retinal coding, spatial (Crick, Marr, & Poggio, 1981) or temporal (Ahissar & Arieli, 2001).

During natural vision, efference copies convey information about eye trajectory (Balslev, Himmelbach, Karnath, Borchers, & Odoj, 2012). The most accurate efference-copy based

Corresponding author:
Ehud Ahissar, Department of Neurobiology, Weizmann Institute of Science, Rehovot, Israel.
Email: ehud.ahissar@weizmann.ac.il

process suggested so far is probably a process proposed to be involved in the control of microsaccades, in which adaptation to perturbations of 15 arcmin were documented (Havermann, Cherici, Rucci, & Lappe, 2014). Such resolution is clearly not sufficient for resolving fine (often <1 arcmin) spatial ambiguities of foveal vision. As one cannot expect efference copies, which are generated at central motor stations, to be accurate at the level of individual motor spikes, a retina-based mechanism that can encode the trajectory of eyeM appears to be instrumental for accurate vision. Importantly, the existence of such a retinal mechanism is indicated also by a study of visual stability (Poletti, Listorti, & Rucci, 2010). As typical trajectories of fixational eyeM (FeyeM) are in the order of several arcminutes (a few tens of foveal receptive fields [RFs]) per each fixational pause (Barlow, 1952; Cherici, Kuang, Poletti, & Rucci, 2012; Eizenman, Hallett, & Frecker, 1985; Engbert, 2006; Ratliff & Riggs, 1950; Yarbus, 1967), a mechanism at a resolution of individual retinal RFs appears to be necessary. We propose here one plausible mechanism, which is based on simple RFs.

According to a widely accepted afferent scheme, cortical simple cells receive inputs from elongated arrays of retinal ganglion cells via parallel channels through the thalamus (Ferster & Miller, 2000; Hubel, 1996; Reid, 2001). These elongated arrays are composed of ganglion cells that share the same polarity, ON or OFF, a segregation that is refined by FeyeM during development (Rucci, Edelman, & Wray, 2000). As a result of this anatomical arrangement, simple cells respond most vigorously to bars of the appropriate polarity that are oriented parallel to the elongated axis of their RF and are flashed on or moved *across* it (Hubel & Wiesel, 1962). However, they do not respond only to oriented bars: Both in cats and primates, simple cells respond also to single dots flashed within their RF (Gur & Snodderly, 1987; Hirsch, Alonso, & Reid, 1995; Hubel & Wiesel, 1962; Tsao, Conway, & Livingstone, 2003). Moreover, simple cells and other cortical cells respond vigorously to single dots or random dot patterns that are moved over their RFs in various directions (Grunewald & Skoumbourdis, 2004; Hubel, 1958; Pack, Born, & Livingstone, 2003; Skottun, Grosof, & De Valois, 1988) in velocities that are typical for FeyeM (Bengi & Thomas, 1968; Riggs, Armington, & Ratliff, 1954). Interestingly, simple cells even show selectivity for dots moving *along* the elongated axis of their RFs (Geisler, Albrecht, Crane, & Stern, 2001). As a result, when complex stimuli are moved over a stationary retina, simple cells respond in a variety of conditions, and are not limited to lines of a specific orientation (Creutzfeldt & Nothdurft, 1978). Crucially, cortical neurons respond reliably also in the reverse, natural case, in which a moving eye scans stationary images during fixation (Snodderly, Kagan, & Gur, 2001).

The size of simple RFs decreases as they get closer to the fovea. However, as recordings approach the fovea, measurements of RF size become difficult due to eyeM, especially in awake primates. Measurements done with the aid of image stabilization, that is, while moving the image according to the on-line-recorded trajectory of the FeyeM, show that at eccentricities of 2° to 9°, the width of subfields of simple RFs in layer four averages around 12 arcmin, and can be as low as 5 arcmin or less (Kagan, Gur, & Snodderly, 2002). We are not aware of stabilized data from smaller eccentricities. Extrapolation of these data suggests that the width of subfields of foveal simple cells is expected to be at the level of a single-cone, which is consistent with a number of anatomical, physiological, and psychophysical indications (Daniel & Whitteridge, 1961; McMahon, Lankheet, Lennie, & Williams, 2000; Polyak, 1941; Smallman, MacLeod, He, & Kentridge, 1996).

Many simple retinal receptive fields (sRFs) contain a row of one polarity (ON or OFF) surrounded by flanks of the opposite polarity (Hubel & Wiesel, 1962; Figure 1). This double-flank type of sRF (“symmetrical flanks,” Table 1 in Hubel & Wiesel, 1962) is capable of

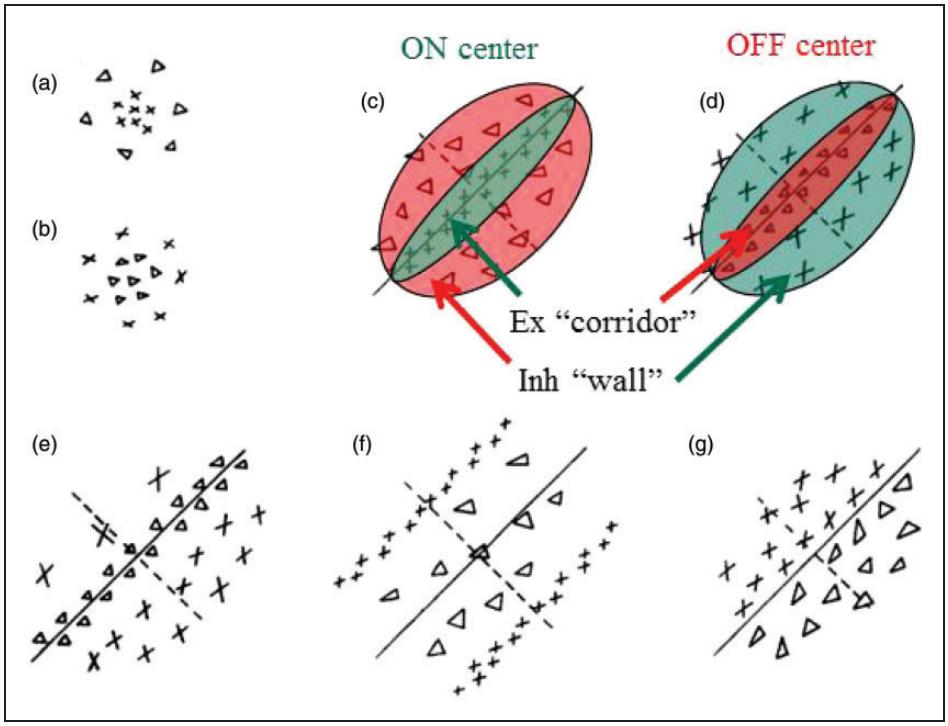

Figure 1. Symmetrical simple receptive fields (ssRFs). Two of the five versions of sRFs presented by Hubel and Wiesel (1962) are of the symmetrical type (colored) which suits eyeM encoding. One is ON-center, and one is OFF-center. In both, the central row ("corridor") is excitatory, and the flanks are inhibitory when the appropriate contrast (dark-to-light or light-to-dark, respectively) is crossed or flashed.

encoding the trajectory of eyeM via the signals generated by the interactions of these movements with the external scene. The special structure of the symmetrical sRFs (ssRFs) produces a kind of an isolated "corridor" through which, at certain retinal velocities, image details are encoded along only one axis during a fixational pause—*along* the elongated axis of the sRF (Figure 2(a), orange arrow). This "corridor protection" against orthogonally moving stimuli (Figure 2(a), black arrow) will hold for retinal velocities inducing a delay of a few milliseconds between the inhibitory and excitatory zones of the RF, a delay during which the inhibition is effective. We tested local drift speed in nine subjects in two experimental settings and found that local drift speed distributes mostly between ~1 and 10 deg/s (Figure 2(b), see experimental details in the legend). These speeds are larger than previous estimations (e.g., 1° deg/s, Cherici et al., 2012) because here we computed the derivative of the raw, non-smoothed eye-position signal, assessing the local speed by which a stimulus crosses the RF. Technically, as the relevant time window for estimating sRF responses is in the order of 1 ms, no smoothing is allowed for data sampled at intervals larger than 1 ms, as is the case here. The down side here is that our data unavoidably contain measurement noise which inflates our distributions. Hence, the upper side of each distribution (Figure 2(b)) should be taken as an upper bound for local drift velocity in that experiment.

Retinal speeds of 1 to 10 deg/s translate to a single-cone delay (i.e., the time a stimulus crosses the RF of one cone) of 8 to 0.8 ms, assuming RF diameter of 0.5 arcmin. Thus, foveal single-cone flanked ssRFs cannot be activated by local drift motion that is perpendicular to

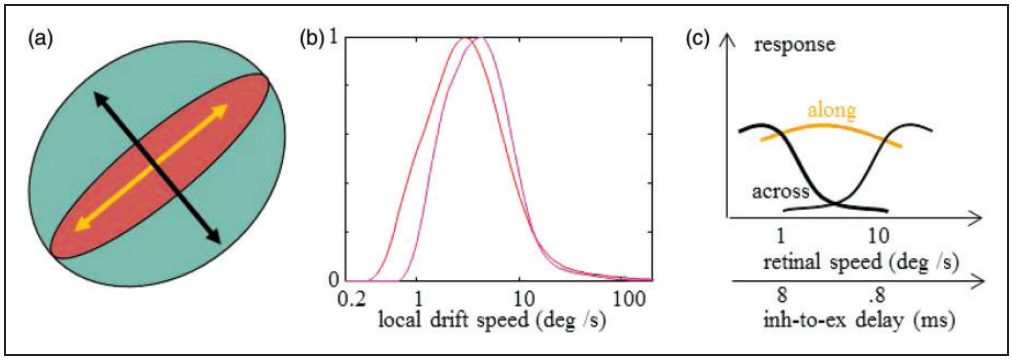

Figure 2. Assumed velocity tuning of ssRF cortical cells. (a) orthogonal movement directions in relation to the orientation of the ssRF. (b) Distribution of local drift speeds computed as the difference between consecutive ocular angles per time difference. Data presented for two experiments: In Experiment 1 (red curve, five subjects, total 797 s, eyeM sampling at 250 Hz, viewing distance 0.45 m) observers fixated on a stationary cross ($0.6^\circ \times 0.6^\circ$) with a dynamic background ($14.8^\circ \times 14.8^\circ$ of a random dot stereogram whose dots fluctuated at 120 Hz); in Experiment 2 (magenta, four subjects, total 315 s, eyeM sampling at 500 Hz, viewing distance 0.6 m) observers viewed at fixation 1-Hz-flashing Gabor patches (3 cyc/deg and envelope of four cycles, with random contrast of seven logarithmic levels in the range 0.8%–50%). Histogram curves were smoothed. Data of Experiment 2 courtesy of Yoram Bonneh. See text for reservations regarding measurement noise. (c) ssRFs are assumed to be tuned for responding to movements *along* their long axis at typical FeyeM speeds (~1–10 deg/s, orange), and to movements *across* their long axis at lower or higher velocities (black). The inh-to-ex delays corresponding to a single-cone flank of 0.5 arcmin are denoted below the speed values.

their long axis, regardless of the stimulus shape, due to the inhibition that leads the excitation by ~1 to 10 ms. In contrast, a movement *along* the long axis should activate the ssRF cell—strongly for small visual elements that only activate the "corridor," and less strongly for larger elements that activate in parallel the flanks; in the latter case the inhibition generated at the flanks will trail the corridor excitation and shorten its response bursts.

The tuning curves for an ssRF cell that scans an image detail *along* versus *across* its long axis are expected to be complementary: with typical FeyeM speeds (~1–10 deg/s) the responses *along* (Figure 2(c), orange) should be stronger than *across* whereas at lower or higher speeds the responses *across* (Figure 2(c), black curves) should be stronger. Consistent with this analysis, a significant portion of cortical simple cells is found to be tuned to either low (often <1 deg/s) or high (often >10 deg/s) retinal speeds of bars moving *across* the sRF (Dow, 1974; Orban, Kennedy, & Maes, 1981; Wilson & Sherman, 1976). Our prediction is thus that these low and high speed tuned cells (for bars moving *across* the RF) possess ssRFs and that they will respond reliably to movements *along* their RF with speeds of ~1–10 deg/s.

The "protected corridor" of ssRFs during FeyeM offers two important functions. The first is the encoding of the trajectory of FeyeM. ssRFs report the transition of contrasts *along* their long axis. With a stationary scene, this transition is caused by FeyeM; each location in the visual field is sampled in succession by neighboring retinal cells such that a single external dot "draws" a contour over the retina, which crosses ssRFs at various directions, determined by the direction of movement of the eye (Figure 3). Each ssRF is "labeled" with a direction of eyeM; it can never be activated during FeyeM if the concurrent eyeM does not include a motion component in its orientation. This information is recoded in the visual cortex by the firing of the corresponding simple cells. With a single stationary external dot, the only ssRF

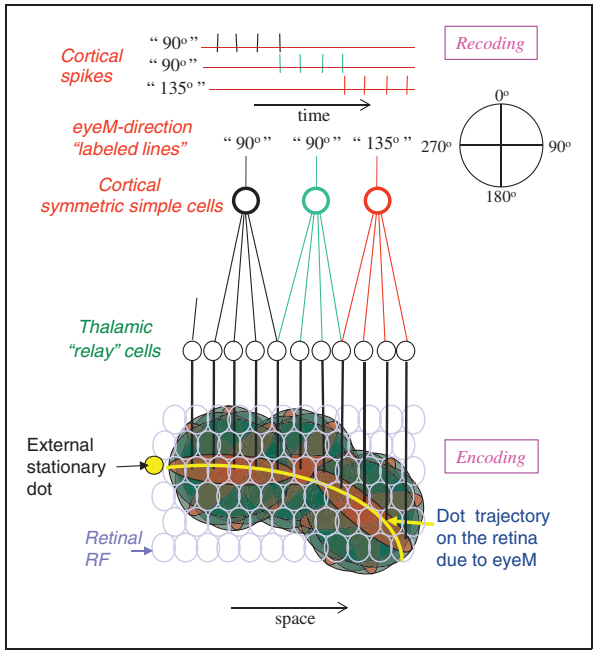

Figure 3. Encoding of eyeM trajectory via ssRFs. Bottom, retinal encoding during eye movement. A schematic description of a retinal trajectory (yellow curve) of a stationary external dot (yellow spot) over a moving retinal mosaic of foveal ganglion RFs. Cortical symmetric simple cells receive their inputs, via thalamocortical neurons, from elongated ssRFs retinal fields. Only the cells whose RF's long axis is aligned with the direction of eyeM (three are shown) can be activated; the rest (cells not shown) remain silent. Top, recoding by cortical symmetric simple cells; each cell codes eyeM direction, by labeled-line coding, and movement speed and distance (in retinal coordinates) by firing rate and count, respectively.

cells that fire are those whose "label" corresponds to the direction of eyeM: Their spike rate indicates the retinal speed and the number of spikes indicates the retinal distance (in no. of retinal RFs) of the concurrent eyeM, in their "labeled" direction. Complex cells would be more reliable coders of eyeM trajectory in such sparse scenes due to their integration of several simple cells of the same orientation. Naturally, in complex stationary scenes a coding scheme that is based on populations of ssRF cells must be applied.

With a nonstationary scene, local motions in the scene are resolved based on relative information among ssRFs (Ahissar & Arieli, 2012). The "mean field" of cortical ssRF cells over the entire visual field codes the motion of the eye, whereas local deviations code local motions in the external scene. With such coding scheme, it is easy to see why various stationary images can evoke an illusory perception of motion. Images such as the Ouchi image (Figure 4(a)) contain arrays with orthogonally orienting bars. When the eye drifts over the image, horizontal ssRFs are activated only when there is a horizontal component of motion and vertical ssRFs are activated only when a vertical component exists. With a uniform image (Figure 4(b)), the cortical coding is uniform over the entire visual field and the interpretation of the visual system is that these activations code eyeM. However, with the Ouchi image, the activation of ssRF cells is not uniform over the field; the different orientations of the bars will induce different activation levels within each family of ssRF cells (e.g., within horizontal cells and within vertical cells) in the center and the background of the image—in the current example, horizontal and vertical eyeM should induce stronger

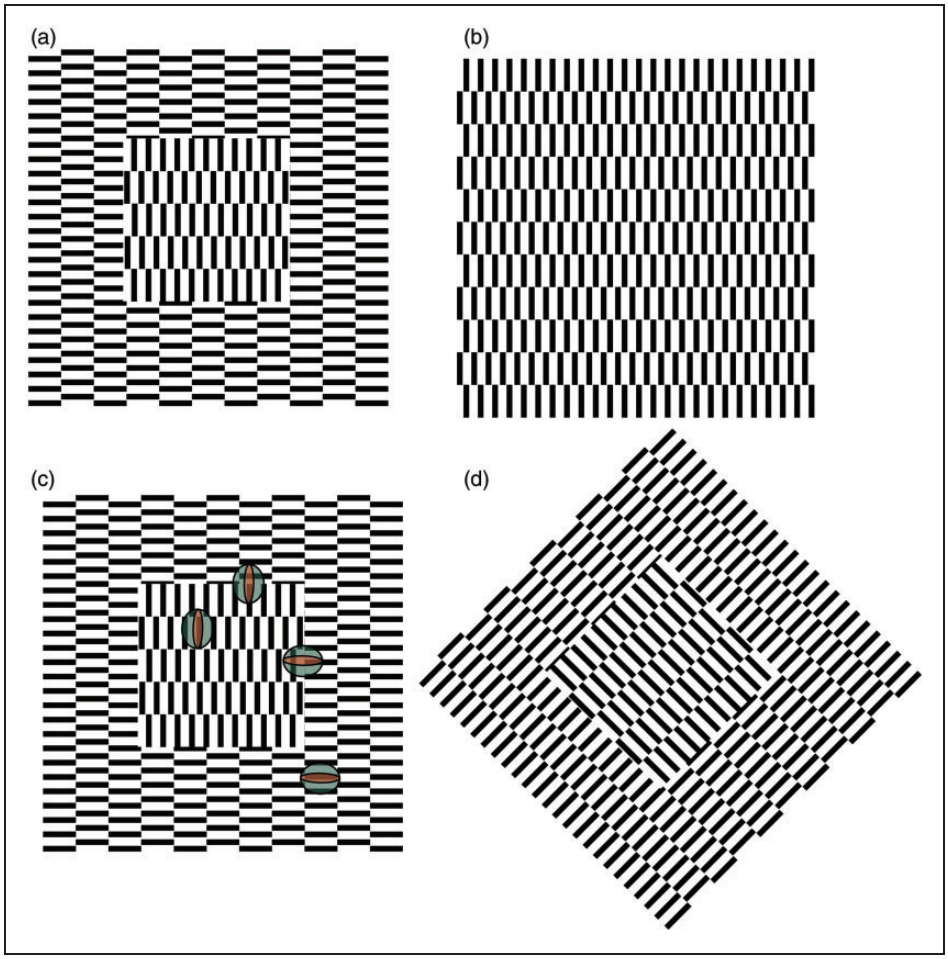

Figure 4. FeyeM-induced motion illusions. (a) Ouchi illusion. The central array of vertical bars appears to jitter around with motion reflecting FeyeM when contrasted with an orthogonal array. (b) A uniform array does not evoke motion illusion. (c) Illustration of relationships between ssRFs and the image. (d) A demonstration of eyeM direction effect. Hold the tip of a pencil near the top corner and then move it down and up over the image while constantly pursuing the tip with your gaze. The center moves to opposite directions with opposite eyeM directions. Now rotate the page 90 degrees and repeat the procedure – directions reverse.

activation in the background and the center, respectively (Figure 4(c)). Upon such conflicts, the visual system appears to interpret one field (center or background) as coding eyeM (and thus stationary external image) and the other as coding external motion. The direction of the perceived external motion depends on the direction of the current eyeM (Figure 4(d)). This ssRF-based explanation can provide a mechanistic basis for computational accounts of the Ouchi illusion or of optical-flow sensing (Fermüller, Pless, & Aloimonos, 2000).

The second function offered by the protected corridor of the ssRFs is the allowance of independent processing of orthogonal axes of the visual image, primarily vertical versus horizontal (''1-D vision''). For example, when the corridor is fully protected by its inhibitory flanks, a vertical eyeM will never activate a horizontal ssRF and a horizontal

eyeM will never activate a vertical ssRF, whatever image is scanned. This is because no intensity change can occur along one axis (say, horizontal) due to a motion along its orthogonal axis (say, vertical). Using the outputs of horizontal versus vertical ssRFs, the visual system can thus process separately orthogonal projections of intensity changes in the image—a kind of Cartesian decomposition of the 2-D image into two 1-D projections, processed sequentially in a ''time sharing'' manner. The horizontal and vertical axes are of special interest in this case because their segregation allows independent control of horizontal and vertical ocular muscles via dedicated motor-sensory-motor loops (Ahissar, Arieli, Fried, & Bonneh, 2014; Saig, Gordon, Assa, Arieli, & Ahissar, 2012; Sherman et al., 2013). Binding of activities in such 1-D perceptual loops to form 2-D percepts may be based on a priori geometrical rules that are imprinted in the system, image-based information such as luminance discontinuities that are coded by synchronous activities (Kuang, Poletti, Victor, & Rucci, 2012; Rucci, 2008) among loops processing different directions and dynamic convergence on pre-learned interloop 2-D attractors.

Acknowledgements

We thank Yoram Bonneh for the data for Figure 2 and for illuminative discussions. E. A. holds the Helen Diller Family Professorial Chair of Neurobiology.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the office of the Chief Scientist, Israeli Ministry of Health, The Minerva Foundation funded by the Federal German Ministry for Education and Research, the Irving B. Harris Fund for New Directions in Brain Research, the Weizmann—UK Joint Research Program, and the Israel Science Foundation (grant 1127/14).

References

Ahissar, E., & Arieli, A. (2001). Figuring space by time. *Neuron*, *32*, 185–201.

Ahissar, E., & Arieli, A. (2012). Seeing via miniature eye movements: A dynamic hypothesis for vision. *Frontiers in Computational Neuroscience*, *6*, 89.

Ahissar, E., Arieli, A., Fried, M., & Bonneh, Y. (2014). On the possible roles of saccades and drifts in visual perception. *Vision Research* [Ebub ahead of print].

Balslev, D., Himmelbach, M., Karnath, H.-O., Borchers, S., & Odoj, B. (2012). Eye proprioception used for visual localization only if in conflict with the oculomotor plan. *The Journal of Neuroscience*, *32*, 8569–8573.

Barlow, H. B. (1952). Eye movements during fixation. *The Journal of Physiology*, *116*, 290–306.

Bengi, H., & Thomas, J. G. (1968). Three electronic methods for recording ocular tremor. *Medical Biology and Engineering*, *6*, 171–179.

Burak, Y., Rokni, U., Meister, M., & Sompolinsky, H. (2010). Bayesian model of dynamic image stabilization in the visual system. *Proceedings of the National Academy of Sciences of the U S A*, *107*, 19525–19530.

Cherici, C., Kuang, X., Poletti, M., & Rucci, M. (2012). Precision of sustained fixation in trained and untrained observers. *Journal of Vision*, *12*, 31.

Creutzfeldt, O. D., & Nothdurft, H. C. (1978). Representation of complex visual stimuli in the brain. *Naturwissenschaften*, *65*, 307–318.

Crick, F. H. C., Marr, D. C., & Poggio, T. (1981). An information-processing approach to understanding the visual cortex. In F. O. Schmitt (Ed.), *The organization of the cerebral cortex* (pp. 505–533). Cambridge, England: Cambridge University Press.

Daniel, P. M., & Whitteridge, D. (1961). The representation of the visual field on the cerebral cortex in monkeys. *The Journal of Physiology (Paris)*, *159*, 203–221.

Dow, B. M. (1974). Functional classes of cells and their laminar distribution in monkey visual cortex. *Journal of Neurophysiology*, *37*, 927–946.

Eizenman, M., Hallett, P. E., & Frecker, R. C. (1985). Power spectra for ocular drift and tremor. *Vision Research*, *25*, 1635–1640.

Engbert, R. (2006). Microsaccades: A microcosm for research on oculomotor control, attention, and visual perception. *Progress in Brain Research*, *154*, 177–192.

Fermüller, C., Pless, R., & Aloimonos, Y. (2000). The Ouchi illusion as an artifact of biased flow estimation. *Vision Research*, *40*, 77–95.

Ferster, D., & Miller, K. D. (2000). Neural mechanisms of orientation selectivity in the visual cortex. *Annual Review of Neuroscience*, *23*, 441–471.

Geisler, W. S., Albrecht, D. G., Crane, A. M., & Stern, L. (2001). Motion direction signals in the primary visual cortex of cat and monkey. *Visual Neuroscience*, *18*, 501–516.

Grunewald, A., & Skoumbourdis, E. K. (2004). The integration of multiple stimulus features by V1 neurons. *The Journal of Neuroscience*, *24*, 9185–9194.

Gur, M., & Snodderly, D. M. (1987). Studying striate cortex neurons in behaving monkeys: Benefits of image stabilization. *Vision Research*, *27*, 2081–2087.

Havermann, K., Cherici, C., Rucci, M., & Lappe, M. (2014). Fine-scale plasticity of microscopic saccades. *The Journal of Neuroscience*, *34*, 11665–11672.

Hirsch, J. A., Alonso, J. M., & Reid, R. C. (1995). Visually evoked calcium action potentials in cat striate cortex. *Nature*, *378*, 612–616.

Hubel, D. (1958). Cortical unit responses to visual stimuli in nonanesthetized cats. *American Journal of Ophthalmology*, *46*, 110–121.

Hubel, D. (1996). A big step along the visual pathway. *Nature*, *380*, 197–198.

Hubel, D. H., & Wiesel, T. N. (1962). Receptive fields, binocular interaction and functional architecture in the cat visual cortex. *The Journal of Physiology*, *160*, 106–154.

Kagan, I., Gur, M., & Snodderly, D. M. (2002). Spatial organization of receptive fields of V1 neurons of alert monkeys: Comparison with responses to gratings. *Journal of Neurophysiology*, *88*, 2557–2574.

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

McMahon, M. J., Lankheet, M. J., Lennie, P., & Williams, D. R. (2000). Fine structure of parvocellular receptive fields in the primate fovea revealed by laser interferometry. *The Journal of Neuroscience*, *20*, 2043–2053.

Orban, G., Kennedy, H., & Maes, H. (1981). Response to movement of neurons in areas 17 and 18 of the cat: Velocity sensitivity. *Journal of Neurophysiology*, *45*, 1043–1058.

Pack, C. C., Born, R. T., & Livingstone, M. S. (2003). Two-dimensional substructure of stereo and motion interactions in macaque visual cortex. *Neuron*, *37*, 525–535.

Poletti, M., Listorti, C., & Rucci, M. (2010). Stability of the visual world during eye drift. *The Journal of Neuroscience*, *30*, 11143–11150.

Polyak, S. L. (1941). *The retina*. Chicago, IL: University of Chicago.

Ratliff, F., & Riggs, L. A. (1950). Involuntary motions of the eye during monocular fixation. *Journal of Experimental Psychology*, *40*(6), 687–701.

Reid, R. C. (2001). Divergence and reconvergence: Multielectrode analysis of feedforward connections in the visual system. *Progress in Brain Research*, *130*, 141–154.

Riggs, L. A., Armington, J. C., & Ratliff, F. (1954). Motions of the retinal image during fixation. *Journal of the Optical Society of America*, *44*, 315–321.

Rucci, M. (2008). Fixational eye movements, natural image statistics, and fine spatial vision. *Network*, *19*, 253–285.

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

Saig, A., Gordon, G., Assa, E., Arieli, A., & Ahissar, E. (2012). Motor-sensory confluence in tactile perception. *The Journal of Neuroscience*, *32*, 14022–14032.

Sherman, D., Oram, T., Deutsch, D., Gordon, G., Ahissar, E., & Harel, D. (2013). Tactile modulation of whisking via the brainstem loop: Statechart modeling and experimental validation. *PLoS One*, *8*, e79831.

Skottun, B. C., Grosof, D. H., & De Valois, R. L. (1988). Responses of simple and complex cells to random dot patterns: A quantitative comparison. *Journal of Neurophysiology*, *59*, 1719–1735.

Smallman, H. S., MacLeod, D. I., He, S., & Kentridge, R. W. (1996). Fine grain of the neural representation of human spatial vision. *The Journal of Neuroscience*, *16*, 1852–1859.

Snodderly, D. M., Kagan, I., & Gur, M. (2001). Selective activation of visual cortex neurons by fixational eye movements: Implications for neural coding. *Visual Neuroscience*, *18*, 259–277.

Tsao, D. Y., Conway, B. R., & Livingstone, M. S. (2003). Receptive fields of disparity-tuned simple cells in macaque V1. *Neuron*, *38*, 103–114.

Wilson, J., & Sherman, S. M. (1976). Receptive-field characteristics of neurons in cat striate cortex: Changes with visual field eccentricity. *Journal of Neurophysiology, 39*, 512–533.

Yarbus, A. L. (1967). *Eye movements and vision*. New York, NY: Plenum.