On texture, form, and fixational eye movements

Tatyana O. Sharpee

Computational Neurobiology Laboratory, Salk Institute for Biological Studies, La Jolla, CA 92037

Abstract:

Recent studies show that small movements of the eye that occur during fixation are controlled in the brain by similar neural mechanisms as large eye movements. Information theory has been successful in explaining many properties of large eye movements. Could it also help us understand the smaller eye movements that are much more difficult to study experimentally? Here I describe new predictions for how small amplitude fixational eye movements should be modulated by visual context in order to improve visual perception. In particular, the amplitude of fixational eye movements is predicted to differ when localizing edges defined by changes in texture or luminance.

Impressionist paintings serve to illustrate the point that visual forms can be defined by changes in textures as well as by changes in brightness [1**, 2-5] (see Fig. 1 for an example). Recent studies have invigorated the debate in visual neuroscience as to whether form or texture are the primary drivers for visual perception [6, 7**, 8]. For example, often shapes defined by textures are perceived more readily than those based on outlines [8]. However, textures themselves are determined by conjunctions of local shape elements, such as the predominance of signals at some orientation or conjunctions of edges [9]. Furthermore, cartoons illustrate that we can perceive shapes based on their outlines alone, without any textural information. Thus, both mechanisms work in parallel to allow for visual perception. In the primary visual cortex (V1), neural responses are tuned to specific combinations of angles at specific positions. Presumably this explains why V1 neurons are better at discriminating individual samples with a shared texture than different texture types from each other [7**]. This situation changes in the secondary visual cortex (V2) where neurons trade the ability to distinguish individual samples for their ability to distinguish between different texture types [6,7**]. This review will first discuss recent results on the neural mechanisms for detecting edges defined in changes in luminance or texture. Then, we will discuss how differences in neural mechanisms translate into different predictions for optimal eye movements based on information theory.

Neural mechanisms for detecting edges defined by textures.

Because textures are defined as patterns with position-invariant statistical properties [10], the responses of neurons tuned to textures are often analyzed using multi-stage models that combine position-invariance with selective tuning to conjunctions of edges of different angles [1**, 2-5,11,12] (Figure 2). Analyses of V2 responses to natural stimuli using such models have yielded three organizing principles for their feature-selectivity [11]. First, the responses of V2 neurons are based on conjunctions of multiple edges at nearby positions. The selectivity to this preferred pattern is strengthened through the cross-orientation suppression where excitatory edge patterns are paired with suppressive edges of approximately orthogonal orientation [11].

Second, there is position invariance in at least two different space scales: at the level of individual edges that locally form the so-called quadrature pairs [13,14], and with respect to position invariance of the whole relevant pattern. This latter type of more global position invariance is the one that would be the most relevant for mediating texture selectivity. Importantly, some V2 neurons used biphasic pooling masks that can be used to detect edges defined by changes in textural characteristics across the boundary. The pooling masks of V2 neurons are computationally equivalent to the receptive fields of V1 neurons applied to luminance gratings (see Fig. 3 for an example). These three properties –cross-orientation suppression, local position invariance through quadrature pairing, and combinations of biphasic/monophasic pooling masks where observed for each of the sub-populations of V2 neurons that were previously identified based on the diversity of their preferred orientation patterns and temporal characteristics [11,15-18]. Thus, V2 neurons have the abilities to signal the presence of different types of textures and to detect edges defined by changes in texture, using similar computational principles that have been applied to V1 responses to decode position of luminance-defined edges.

Small fixational eye movements as a maximally informative information gathering strategy. This brings us to the question of how small magnitude eye movements that occur during fixation [19**, 20, 21**] could help identify object boundaries. [The effects that we will discuss should work in addition to the mechanisms that remove the redundancy present in natural scenes and which are covered in the recent review [21**].] From the information-theoretic point of view, there are two options for improving estimation accuracy. One option is to keep the eye position as steady as possible to allow the integration of neural responses across time to obtain greater accuracy. Ignoring the case of full image stabilization that causes images to fade, there is evidence of volitional control for microsaccade amplitude and frequency [22-24]. The second option is to actively move the eye to such location that would maximally reduce the estimation uncertainty with respect to object position and/or shape. Here also there is evidence that eye movement position can be controlled to the accuracy of a few arcmin when performing a virtual task of threading a needle [25**]. The trade-off between these two strategies depends on the visual input statistics, as well as the noise characteristics of individual neural responses and number of neurons participating in the sensory estimation task. Two lines of theoretical work inform the answer. First, we can build on the analysis of maximally informative search strategies that aim to localize odorant-emitting target location based on chemosensory cues [26-28]. In that problem, the authors showed that in regime of low odorant concentrations, which also correspond to turbulent flows, it is better to perform active searching across varied locations based on single measurements than to wait to achieve a high precision measurement of the gradient at each location. Similar information-theoretic analysis applied to visual search based on large-amplitude eye movements (termed saccades) reproduced several patterns of human eye search behavior, including the inhibition of return [29]. These results are relevant to our discussion of optimal fixational eye movements, because the statistics of natural visual stimuli exhibits is non-Gaussian [30,31], just as in the case of turbulent olfactory flows. From this perspective, one would expect that active sampling across positions should be beneficial compared to averaging neural responses in time.

Distribution of maximally informative eye movements.

What should be the optimal distribution of eye movements to best localize the position of an edge as defined by either luminance or textural differences? In either case, we would expect that there will be certain directions that would be maximally informative about the object shape [32]. Indeed, for some types of larger fixational eye movements, there is evidence that their properties depend on local visual context [19,21]. For determining boundary location, eye movements aligned perpendicularly with the tentatively detected object boundary would be most informative. Here we will discuss optimal distribution of eye movements, such as eye tremor [20], along that direction. Consider an edge defined by either a change in luminance or texture. The input strength to a neuron with a receptive field (RF) shown in Figure 4, panel A varies with edge position according to a curve shown in panel B. To maximize the mutual information [33] transmitted about the edge location, one should match the distribution of sampled locations to the input-output function of the neural (or more generally neural population) tasked with coding the edge position. Typically, in sensory systems, the input distribution is fixed while the width of the input-output function adaptively changes to maximize the information transmitted [34-36]. However, the eye movements make it possible to control the input distribution. Following [26], we can ask what shifts in the eye position would yield the greatest reduction in the uncertainty associated with the edge position. In Figure 4 we show the results for assuming two different nonlinearities associated with the probability of a spike as a function of input projection on the neuron's RF. Results in panels C, D pertain to the case of a sharp nonlinearity as a function of edge position. This case corresponds to either responses of single neurons with low input noise or the effective response of a summed output from large neural populations. Results depicts in panels E and F describe the case where the response changes more gradually as a function of edge position. In both cases, one observes that the width of the distribution describing the expected change in entropy [33], a measure of uncertainty regarding edge position, scales with the width of the effective neural nonlinearity.

These results lead to a number of predictions for the predicted distribution of small eye movements that occur during individual fixations. For example, one would expect to find differences in eye movements during tasks involving the detection of edges defined by changes in luminance vs. texture. The reason is that the number of neurons in V2 that are tuned to edges in texture is ~25% [11]. Although based on a relatively small dataset of 80 neurons, the corresponding number in V1 for neurons tuned to edges of certain orientation would be close to 100% [37]. Thus, at the behavioral level one expects to be able to pool across large number of neurons when analyzing the position of edges defined by changes in luminance vs. texture. Second, as the sharpness of neuronal nonlinearities is affected by contrast, one can expect larger eye movements when localizing edges in the presence of visual clutter or at low light intensities.

Ultimately, if these predictions are verified, the information-theoretic framework will make it possible to use eye movements as a tool for inferring the number of neurons contributing to different types of visual object discriminations tasks. Given that eye movement statistics is altered in subjects with autism [38], attention deficit [39] or other neurological disorders [40], analyzing these statistics within information theory increase the specificity of diagnoses delivered based on eye movements.

Figure 1. (**A**) Example image with boundaries defined by either changes in luminance or textures. (**B**) A set of relevant edge features for an example V2 neuron. Data from [18] re-analyzed using the three-stage position-invariant model [11] (see also Figure 2). This neuron 'e0043' was identified as belonging to the sub-population with relatively homogeneous feature selectivity across space. Blue and red denote excitatory and suppressive features, respectively; opacity is proportional to the weight with which this feature affects the neural spike probability. (**C**) Example V2 neuron ('e109') from the sub-population with heterogeneous tuning across space.

Figure 2. A three-stage model for characterizing responses of neurons selective to textures. The model incorporates selectivity for multiple excitatory and suppressive components at each position. This operation is repeated across space (red, green, and blue channels). Within each channel, the stimulus patch is projected onto a set of relevant features (same for all patches and shown here as heat maps) to which we refer as first-order features. The output of a projection onto a given feature is passed through a quadratic function (with a potentially non-zero linear term) [1**]. These outputs are summed and passed through a compressive nonlinearity. This part of the model is designed to describe heterogeneous center-surround interactions, because the number of features and their spatial arrangement is not pre-specified and includes both excitatory and suppressive features (marked with + and – near the arrows in the schematic of the block). The output of each quadratic block within each patch is summed, with weights that could be either positive or negative, and the result passed through a soft threshold function to yield a prediction for the firing rate. The block output filtering allows one to connect with filter-rectify-filter (FRF) models [1**,3,4,41-44]. (**B**) Left: Prototypical arrangement of features in the FRF model. Each ellipse denotes a Gabor feature, excitatory (blue) and suppressive (red). Right: equivalent representation in the composite model with a single first-stage filter (blue contour) and a broader block-output filter (dashed line) that includes both positive (+) and negative (-) weights. (**C**) Left: Arrangement of features that can model selectivity to a texture boundary or selectivity to pairs of orientation in the presence of position invariance. Right: Equivalent representation in the composite model with two first-stage filters (blue contours) and an approximately uniform block output filter (denoted by the dashed line). (**D**) Left: Generalization of a FRF model from B that includes cross-orientation suppression between features. The equivalent representation in terms of the composite model has two first stage filters (excitatory in blue and suppressive in red) followed by a biphasic block output filter (dashed line).

Figure 3: Examples of space pooling masks from the third stage of the model for three example V2 neurons. (A) Example V2 neuron with approximately uniform pooling across spatial positions. Such neurons were encountered in ~75% of cases [11]. (B,C) Example biphasic pooling mask for two neurons that could mediate selectivity to texture-defined boundaries. Data from [18] re-analyzed using three-stage position-invariant model [11]. Neurons are 'e0018',

Figure 4: Analysis of maximally informative eye movements for edge localization. (**A**) Example orientation-selective receptive field. (**B**) Relevant stimulus component as a function of edge position for such a neuron. (**C**) Example predicted spike rate for a neuron with a sharp nonlinearity applied to input from panel (B) as a function of edge position. (**D**) Expected reduction in uncertainty in edge location for directed eye positions at

different distances from the edge position. Panels E and F are the same as C and D but for a neuron with a less steep nonlinearity. Nonlinearity could also represent summed output from large neural populations, in this case (C,D) correspond to larger populations than (E, F).

Acknowledgements

The author thanks Margot Larroche, Ryan Rowekamp, Lawrence Sincich, and Sebastien Tawa for discussions. This research was supported by the National Science Foundation (NSF) CAREER award number IIS-1254123 and award number IOS-1556388, the National Eye Institute of the National Institutes of Health under Award Numbers P30EY019005 and a seed grant from the Kavli Institute for Brain and Mind.

References

1**. Schmid AM, Victor JD: **Possible functions of contextual modulations and receptive field nonlinearities: pop-out and texture segmentation**. *Vision Res* 2014, **104**:57-67.
A discussion of possible neural mechanisms for detecting edges defined by texture.

2. Schmid AM, Purpura KP, Victor JD: **Responses to orientation discontinuities in V1 and V2: physiological dissociations and functional implications**. *J Neurosci* 2014, **34**:3559-3578.

3. Landy MS, Graham N: **Visual perception of texture**. In *The visual neurosciences*. Edited by Chalupa LM, Werner S: MIT; 2004:1106-1118.

4. Li G, Yao Z, Wang Z, Yuan N, Talebi V, Tan J, Wang Y, Zhou Y, Baker CL, Jr.: **Form-cue invariant second-order neuronal responses to contrast modulation in primate area V2**. *J Neurosci* 2014, **34**:12081-12092.

5. Zavitz E, Baker CL, Jr.: **Higher order image structure enables boundary segmentation in the absence of luminance or contrast cues**. *J Vis* 2014, **14**.

6. Freeman J, Ziemba CM, Heeger DJ, Simoncelli EP, Movshon JA: **A functional and perceptual signature of the second visual area in primates**. *Nat Neurosci* 2013, **16**:974-981.

7**. Ziemba CM, Freeman J, Movshon JA, Simoncelli EP: **Selectivity and tolerance for visual texture in macaque V2**. *Proc Natl Acad Sci U S A* 2016, **113**:E3140-3149.
This study shows that V2 neurons are better at distinguishing texture types than V2 neurons. In contrast, V1 neurons are better at distinguishing different exemplars of the same texture type than V2 neurons

8. Lettvin JY: **On seeing sidelong**. *The Sciences* 1976, **16**:10-20.

9. Portilla J, Simoncelli E: **A Parametric Texture Model Based on Joint Statistics of ComplexWavelet Coefficients**. *International Journal of Computer Vision* 2000, **40**:49-71.

10. Cross GR, Jain AK: **Markow Random Field Texture Models**. *TransPAMI* 1983, **5**:25-39.

11. Rowekamp RJ, Sharpee TO: **Cross-Orientation Suppression in Visual Area V2**. *Nature Communications* 2017:in press.

12. Yamins DL, DiCarlo JJ: **Using goal-driven deep learning models to understand sensory cortex**. *Nat Neurosci* 2016, **19**:356-365.

13. Movshon JA, Thompson ID, Tolhurst DJ: **Receptive field organization of complex cells in the cat's striate cortex**. *J. Physiol.* 1978, **283**:79-99.

14. Adelson EH, Bergen JR: **Spatiotemporal energy models for the perception of motion**. *J. Opt. Soc. Am. A..* 1985, **2**:284-299.

15. Anzai A, Peng X, Van Essen DC: **Neurons in monkey visual area V2 encode combinations of orientations**. *Nat Neurosci* 2007, **10**:1313-1321.

16. Schmid AM, Purpura KP, Ohiorhenuan IE, Mechler F, Victor JD: **Subpopulations of neurons in visual area V2 perform differentiation and integration operations in space and time**. *Front Syst Neurosci* 2009, **3**:15.

17. Liu L, She L, Chen M, Liu T, Lu HD, Dan Y, Poo MM: **Spatial structure of neuronal receptive field in awake monkey secondary visual cortex (V2)**. *Proc Natl Acad Sci U S A* 2016, **113**:1913-1918.

18. Willmore BD, Prenger RJ, Gallant JL: **Neural representation of natural images in visual area V2**. *J Neurosci* 2010, **30**:2102-2114.

19**. Krauzlis RJ, Goffart L, Hafed ZM: **Neuronal control of fixation and fixational eye movements**. *Philos Trans R Soc Lond B Biol Sci* 2017, **372**.

A recent comprehensive review describing that the same neural mechanisms control fixational eye movements and large amplitude saccades.

20. Martinez-Conde S, Macknik SL, Hubel DH: **The role of fixational eye movements in visual perception**. *Nat Rev Neurosci* 2004, **5**:229-240.

21**. Rucci M, Victor JD: **The unsteady eye: an information-processing stage, not a bug**. *Trends Neurosci* 2015, **38**:195-206.

A recent review of the impact of fixational eye movements on the spatiotemporal statistics of input to the visual system.

22. Hafed ZM, Lovejoy LP, Krauzlis RJ: **Modulation of microsaccades in monkey during a covert visual attention task**. *J Neurosci* 2011, **31**:15219-15230.

23. Pastukhov A, Braun J: **Rare but precious: microsaccades are highly informative about attentional allocation**. *Vision Res* 2010, **50**:1173-1184.

24. Steinman RM, Cunitz RJ, Timberlake GT, Herman M: **Voluntary control of microsaccades during maintained monocular fixation**. *Science* 1967, **155**:1577-1579.

25**. Ko HK, Poletti M, Rucci M: **Microsaccades precisely relocate gaze in a high visual acuity task**. *Nat Neurosci* 2010, **13**:1549-1553.

A demonstration that eye position can be precisely controlled in a virtual 'needle-threading' tasks with gaze alterating between the target and the thread.

26. Vergassola M, Villermaux E, Shraiman BI: **'Infotaxis' as a strategy for searching without gradients**. *Nature* 2007, **445**:406-409.

27. Calhoun AJ, Chalasani SH, Sharpee TO: **Maximally informative foraging by Caenorhabditis elegans**. *Elife* 2014, **3**.

28. Sharpee TO, Calhoun AJ, Chalasani SH: **Information theory of adaptation in neurons, behavior, and mood**. *Curr Opin Neurobiol* 2014, **25**:47-53.

29. Najemnik J, Geisler WS: **Optimal eye movement strategies in visual search**. *Nature* 2005, **434**:387-391.

30. Simoncelli EP, Olshausen BA: **Natural image statistics and neural representation**. *Annu Rev Neurosci* 2001, **24**:1193-1216.

31. Ruderman DL, Bialek W: **Statistics of natural images: Scaling in the woods**. *Phys. Rev. Let.* 1994, **73**:814-817.

32. Li A, Zaidi Q: **Three-dimensional shape from non-homogeneous textures: carved and stretched surfaces**. *J Vis* 2004, **4**:860-878.

33. Cover TM, Thomas JA: *Elements of Information Theory*. New York: Wiley-Interscience; 1991.

34. Wark B, Lundstrom BN, Fairhall A: **Sensory adaptation**. *Curr Opin Neurobiol* 2007, **17**:423-429.

35. Fairhall AL, Lewen GD, Bialek W, de Ruyter Van Steveninck RR: **Efficiency and ambiguity in an adaptive neural code**. *Nature* 2001, **412**:787-792.

36. Maravall M, Petersen RS, Fairhall AL, Arabzadeh E, Diamond ME: **Shifts in coding properties and maintenance of information transmission during adaptation in barrel cortex**. *PLoS Biol* 2007, **5**:e19.

37. Olshausen BA, Field DJ: **How close are we to understanding v1?** *Neural Comput* 2005, **17**:1665-1699.

38. Schmitt LM, Cook EH, Sweeney JA, Mosconi MW: **Saccadic eye movement abnormalities in autism spectrum disorder indicate dysfunctions in cerebellum and brainstem**. *Mol Autism* 2014, **5**:47.

39. Bittencourt J, Velasques B, Teixeira S, Basile LF, Salles JI, Nardi AE, Budde H, Cagy M, Piedade R, Ribeiro P: **Saccadic eye movement applications for psychiatric disorders**. *Neuropsychiatr Dis Treat* 2013, **9**:1393-1409.

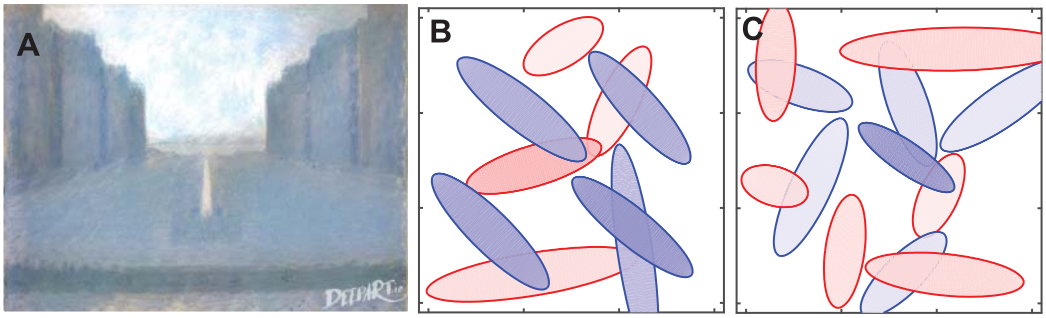
40. Paolozza A, Munn R, Munoz DP, Reynolds JN: **Eye movements reveal sexually dimorphic deficits in children with fetal alcohol spectrum disorder**. *Frontiers in Neuroscience* 2015, **9**:76.

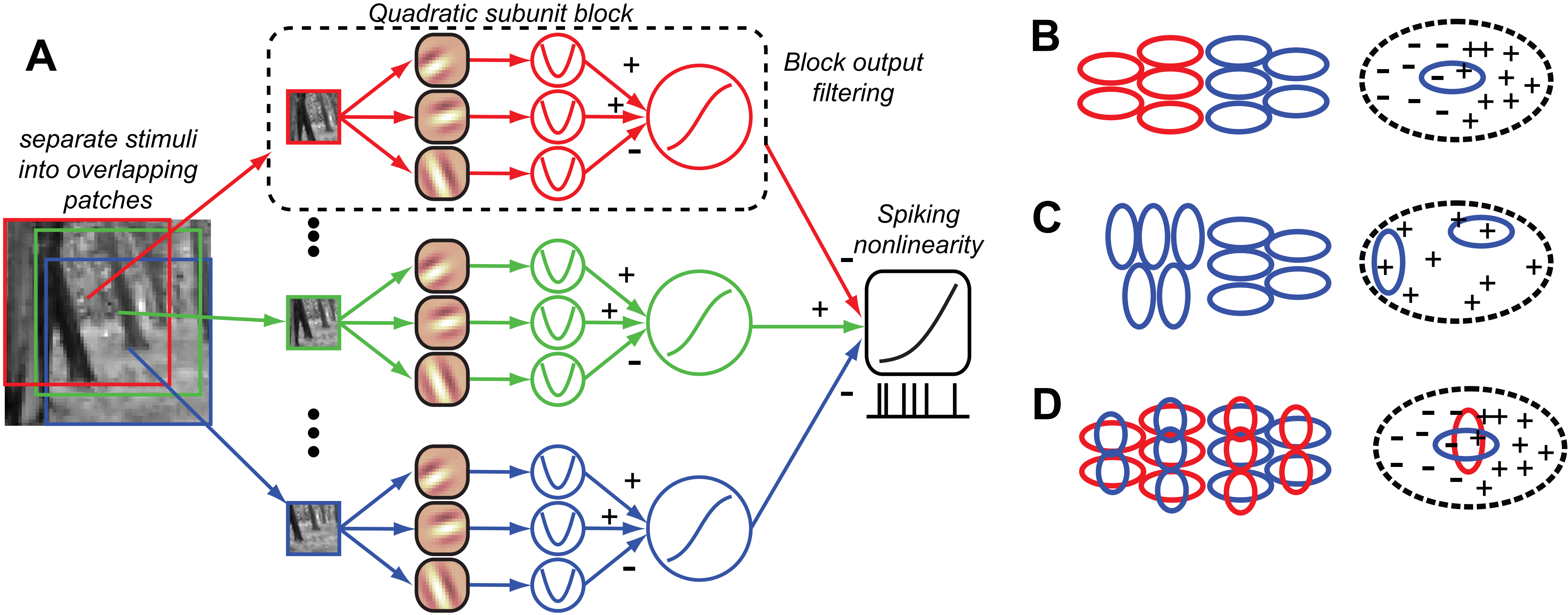
41. Wilson HR, Ferrera VP, Yo C: **A psychophysically motivated model for two-dimensional motion perception**. *Vis Neurosci* 1992, **9**:79-97.

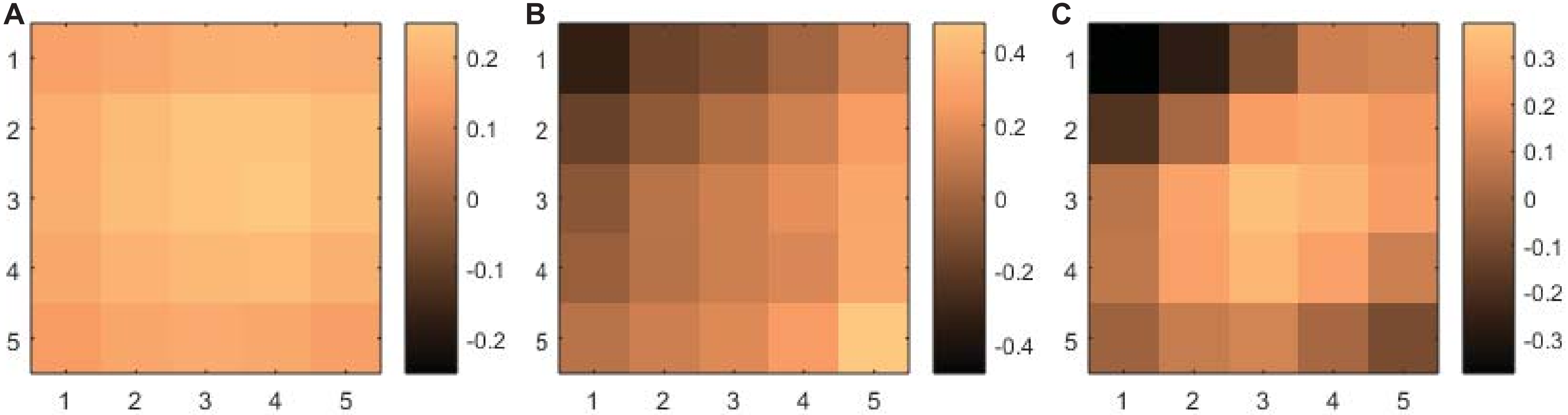
42. Ellemberg D, Allen HA, Hess RF: **Second-order spatial frequency and orientation channels in human vision**. *Vision Res* 2006, **46**:2798-2803.

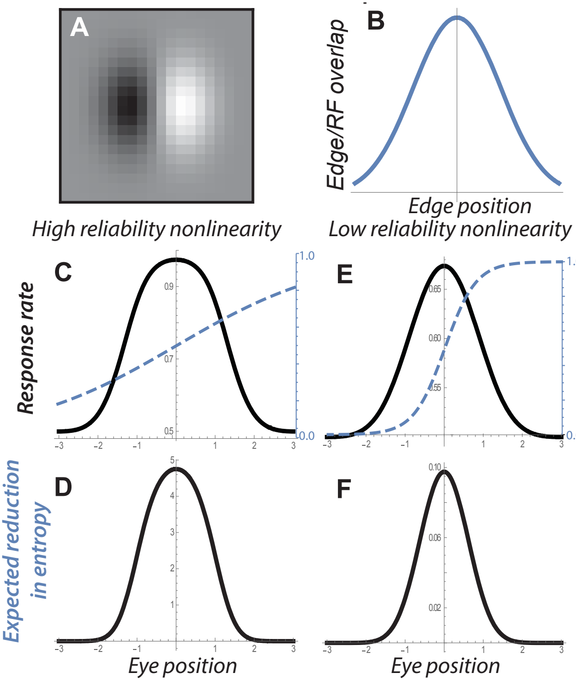
43. Chubb C, Sperling G: **Drift-balanced random stimuli: a general basis for studying non-Fourier motion perception**. *J Opt Soc Am A* 1988, **5**:1986-2007.

44. Zavitz E, Baker CL: **Texture sparseness, but not local phase structure, impairs second-order segmentation**. *Vision Res* 2013, **91**:45-55.


A
DEEPART
B
C


A
Quadratic subunit block
Block output filtering
separate stimuli into overlapping patches
Spiking nonlinearity
+
+
–
+
–
–
B
C
D


A
B
C
1
2
3
4
5
1
2
3
4
5
0.2
0.1
0
-0.1
-0.2
0.4
0.2
0
-0.2
-0.4
0.3
0.2
0.1
0
-0.1
-0.2
-0.3


A
B
Edge/RF overlap
Edge position
High reliability nonlinearity
Low reliability nonlinearity
C
Response rate
0.9
0.7
0.5
1.0
0.0
-3
-2
-1
1
2
3
E
0.65
0.60
0.55
1.0
0.0
-3
-2
-1
1
2
3
D
Expected reduction in entropy
5
4
3
2
1
-3
-2
-1
1
2
3
Eye position
F
0.10
0.06
0.02
-3
-2
-1
1
2
3
Eye position