

Specific aims

One of the primary goals in neuroscience is to figure out simple principles that explain how nervous systems are organized. One of the most successful theories in this vein states that sensory systems should remove redundancies in their inputs to optimize the information they process⁵. This *efficient coding hypothesis* provides us with a mathematical framework to understand how neurons *should* encode information, which can then be experimentally tested against how neurons *actually* do so. Over the past 60 years, efficient coding has successfully explained many experimental findings in different sensory modalities such as vision^{4, 10, 15, 24, 25, 28}, audition²⁰ and touch²³. This hypothesis has been especially successful in the retina, where it can explain many features of retinal encoding such as center-surround receptive fields and ON-OFF pathways^{2-4, 10, 15}. However, we still lack efficient coding predictions for how the retina processes many complex features of the visual world such as color and motion. My work will tackle this problem by providing a theoretical account of how the retina integrates redundant inputs across different color channels (Aim 1) and across time (Aim 2). These results extend our understanding of how retinal physiology can be explained by efficient coding principles.

Aim 1: Expand efficient coding models to encompass chromatic information

Color is a crucial aspect of how we perceive the visual world. Encoding colors starts at the level of cone photoreceptors in the fovea of the retina, which come in three types –Long (L), Medium (M), and Short (S), roughly encoding red, green and blue stimuli, respectively⁸. This color information is ultimately encoded by RGCs, with different types integrating cone inputs differently. Most RGCs (~90% in the fovea and ~45% in the periphery) are midget cells⁹, which, in the fovea, are ‘red-green opponent’ and encode a contrast between L and M inputs⁷⁻⁹. In the periphery, these midget cells are not red-green opponent but instead sum L and M cone inputs^{7, 9, 22}. Why midget cells use different coding strategies for the fovea and for the periphery is still unclear. We hypothesize those different strategies have to do with the ratio between RGCs and cones in the fovea versus the periphery. While the periphery has more cones than RGCs, the opposite is true for the primate fovea, with approximately 3 RGCs for every cone³⁰. Here we will test that hypothesis by building an efficient coding model for chromatic natural images. We will try to replicate the differences in receptive fields between the fovea and the periphery by building efficient coding models with different RGC-cone ratios. We will also test whether we can replicate receptive fields of RGC types other than midget cells, such as parasol and bistratified cells. Completion of this aim will grant us a theoretical understanding of how the retina processes chromatic information.

Aim 2: Expand efficient coding models to explain motion-selectivity in RGCs

Visual scenes are typically in motion, either because of objects moving or optic flow from our own movements. The encoding of visual motion starts as early as the retina, with several subtypes of retinal ganglion cells (RGCs) having stronger responses to one direction of motion than to its opposite^{26, 29}. The major type of motion encoding RGCs are ON-OFF direction-selective ganglion cells³¹, which are divided into four subtypes that respond preferentially to each of the four cardinal directions²⁹. Yet it is still not known whether or how efficient coding principles can explain this finding. My lab previously studied efficient coding in spatiotemporal receptive fields¹², but this study made strong assumptions (independently processed spatial and temporal information) that preclude encoding visual motion. To test whether we can explain motion-selectivity in RGCs from efficient coding principles, I will build an efficient coding model that estimates receptive fields in which space and time are processed together. I predict the efficient coding model will be able to replicate direction selectivity in RGCs, and that we will find four subtypes that encode motion in the four cardinal directions²⁹. Completion of this aim will result in testable predictions for efficient coding of motion in natural images, which we will be able to compare to experimental data from the Field Lab at UCLA.

Significance

There are three explanatory processes that underlie scientific explanation: Descriptive (what), mechanistic (how) and normative (why) explanations¹⁹. The retina is fairly well-understood both from

descriptive and mechanistic perspectives^{9, 14}. However, normative models that explain the purpose of the system are still lacking. Efficient coding is one of the most successful theories that can explain how the retina is organized, but what it can explain is still relatively sparse compared to what is mechanistically known. The current work will try to expand efficient coding theory to replicate how the retina processes color and motion.

Mechanistic explanations of retinal processing

Retinal processing of visual information follows a well-known structure^{9, 14}. First, photoreceptors transform light from the outside world into electrical activity. They then send this information to bipolar cells, and bipolar cells send this information to retinal ganglion cells (RGCs). These RGCs are the output layer of the retina. Their axons form the optic nerve and send information to the thalamus, which then transmits it to the primary visual cortex. RGCs are separated into two different pathways (ON and OFF), and each neuron within a pathway processes a small region of visual space — its receptive field. These receptive fields form ‘mosaics’ (one per RGC type) that tile visual space. The receptive fields of RGCs have a center-surround organization: ON RGCs encode light in the center and dark in the surround, and vice-versa for OFF RGCs.

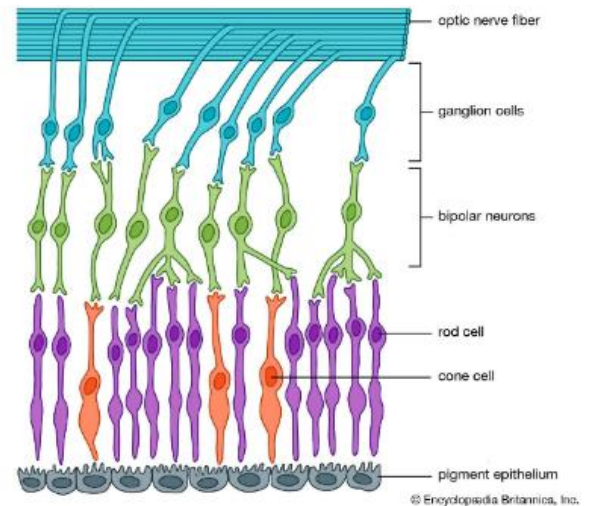


Figure 1: Schematic structure of the retina

Normative explanations of retinal processing

Early theoretical work on efficient coding explained how this center-surround organization arises from decorrelation, both for achromatic and for color inputs^{2, 3}. However, these made many simplifying assumptions, including an infinite number of neurons and linear output responses. While these assumptions help make the problem mathematically tractable, such models neglect key biological realities like strictly positive neural firing rates. More recent work has leveraged machine learning to make efficient coding models that incorporate non-linear output responses and a limited number of neurons^{12, 13, 15, 21, 24}, accounting for the separation of RGCs into different subtypes, with neurons within a subtype forming a mosaic. However, the inputs to RGCs are much more complex than static achromatic images – RGCs receive input images from multiple color channels, and visual scenes are usually in motion. While efficient coding can predict how the retina should process achromatic stimuli, its predictions for color and motion processing – two crucial aspects of natural stimuli – involve simplifying assumptions (e.g. linearity¹ and space-time separability¹²) which are difficult to relate to retinal physiology^{7, 16}.

Innovation

Technical innovation: To complete either aims, I will need to develop new machine learning techniques to train efficient coding models with multiple correlated channels (cones or latencies), which implies increasing the number of parameters by multiple folds. I will solve this overparameterizing problem by designing new methods to parametrize receptive fields across color channels and latencies. By doing so, we will pave the way for future research to solve efficient coding problems with very larger number of parameters.

Conceptual innovation: Most of the efficient coding research in vision involves a single input channel that is encoded by a large number of neurons. However, neurons in the retina have multiple correlated input channels, such as different colors and latencies. This project is conceptually innovative because I consider efficient coding models with multiple correlated channels. By doing so, we will learn how

efficient coding models handle correlated channels, and whether this solution is similar to the computations RGCs perform.

Approach

Previous work

Our lab previously built efficient coding models to explain the organization of RGCs. why retinal ON and OFF mosaics are anti-aligned. We also used a similar model to explain why RGCs encode either high spatial and low temporal frequencies, or low spatial and high temporal frequencies. In such models, the input natural images X are multiplied by a weight matrix W , which is then followed by an output nonlinearity. The output is the model RGCs responses r_j for each neuron j . Input and output noise are represented by independent Gaussian distributions with different standard deviations (σ_{in} for the inputs and σ_{out} for the outputs). The weights are optimized to maximize the mutual information between the natural images and the responses of the model RGCs. To represent the biological cost of firing spikes, each neuron is constrained to have the same average firing rate.

Such models not only replicate the center-surround organization, but also their spatial arrangement. Model RGCs are separated into different ON and OFF pathways that process light and dark information, respectively. Each neuron processes a small region of visual space, and each pathway has neurons that are spatially organized to form a 'mosaic' that tiles the entire visual scene. Our lab found that whether efficient coding predicted that these ON and OFF mosaics should be aligned or anti-aligned depends on the noise, with mosaics going from aligned to anti-aligned as both the input and output noise levels increase. Next, I will extend this model to integrate inputs from multiple correlated channels, which will allow us to study how RGCs process color (Aim 1) and motion (Aim 2).

Aim 1: Expand efficient coding models to encompass chromatic information

Background and rationale

Color vision starts with cone photoreceptors, which encode three different wavelengths of light. Short cones (S) encode blue, Medium (M) cones encode green, and Long (L) cones encode red. The three main types of RGCs (Parasol, midget and bistratified cells) each integrate this chromatic information in different ways (Figure 2). Parasol cells are achromatic and integrate cones independently of their type. Midget cells integrate L and M cones, and how they do so vary from the fovea to the periphery. In the fovea, midget cells receive excitatory inputs from a single cone in the receptive field center (either L or M), and inhibitory inputs from both L and M cones in the surround. In the periphery, both the center and the surround of midget cells encode a sum of L and M cones, with the center encoding the opposite polarity of the surround (i.e. ON-center and OFF-surround or vice-versa). Bistratified cells encode blue: They only have a receptive field center, which encodes ON S inputs and L + M OFF inputs. While we know a lot about how the retina processes chromatic information, what efficient coding models can explain about color vision in the retina is relatively scarce. For example, mathematically tractable models that can explain why color-opponency is efficient are limited to two cones (L and M) instead of three¹. Independent-component analysis can explain why color-opponency with three cones is efficient¹⁸, but this results in receptive fields that are qualitatively different and do not have a center-surround organization. We are still missing efficient coding models that can explain in more details how chromatic inputs are processed

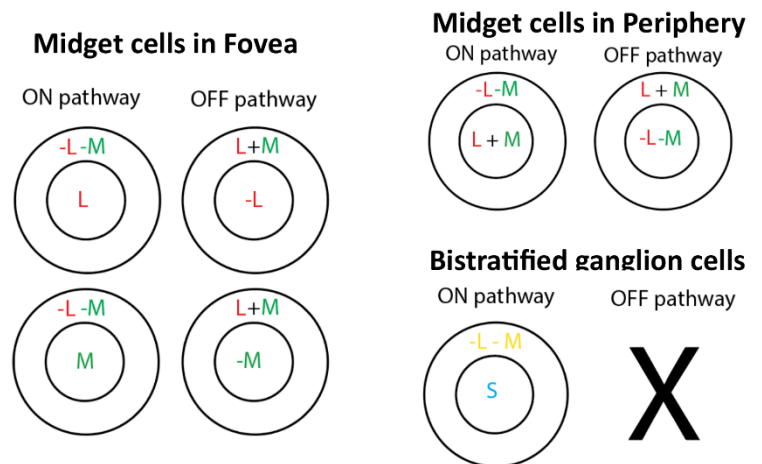


Figure 2: Receptive fields of midget and bistratified ganglion cells.

differently by midget, parasol and bistratified cells. Aim 1 will address this question by building efficient coding models that encode chromatic natural images.

Experimental methods and design

To study how to efficiently encode chromatic natural images, we will use the Kyoto Natural Images Dataset¹¹. This dataset consists of 62 images of natural scenes that were gamma-corrected and transformed from RGB to LMS. The transformation matrix from RGB into LMS responses was estimated by minimizing the prediction error of cone spectral sensitivities¹¹. These responses were further transformed with an empirical cone non-linear function⁶. Each of these images have 1000x1280x3 pixels, which we will convert into training samples of 16x16x3 pixels.

The efficient coding model we use will be similar to other work from our lab^{12, 13}. For each of our model RGCs, we will multiply these chromatic natural images by a linear filter w_j that represents the receptive field of this neuron. The linear filters will have three different channels that process the estimated responses of L, M and S cones (Figure 3). To prevent negative firing rates, the output responses of the filters will be rectified by a parametrized ReLU. The linear filters will be estimated by maximizing the following function¹⁵, which represents the mutual information between the natural image inputs and the model RGCs responses:

$$\text{maximize} : \mathbb{E}_x \log \frac{\det \left(\mathbf{G} \mathbf{W}^T (\mathbf{C}_x + \mathbf{C}_{n_{in}}) \mathbf{W} \mathbf{G} + \mathbf{C}_{n_{out}} \right)}{\det \left(\mathbf{G} \mathbf{W}^T \mathbf{C}_{n_{in}} \mathbf{W} \mathbf{G} + \mathbf{C}_{n_{out}} \right)},$$

Where C_x represents the covariance matrix of the input natural images, $C_{n_{in}}$ and $C_{n_{out}}$ are diagonal matrices that represent the covariance of the input and output noise, respectively, and \mathbf{W} is the weight matrix. \mathbf{G} is a diagonal matrix that represents the local derivatives of the output responses for a specific set of input images. Since the output nonlinearity is a ReLU, the diagonal of \mathbf{G} has binary values (1 if the neuron is firing and 0 if not). This function will be maximized using Adam optimization¹⁷. To represent the metabolic cost of firing spikes, each neuron will be restricted to have a fixed average firing rate.

Having multiple input channels drastically increases the number of parameters, which can make the model significantly harder to optimize. However, we know the filters converge to difference-of-gaussians, which we can use to our advantage. To facilitate training, I will parametrize the linear filter of each neuron to be a difference-of-gaussians:

$$\mathbf{W} = d(e^{-az} - c(e^{-bz}))$$

Where a , b , c and d are vectors of length three that represent each of the three color channels. z is the distance from the receptive field center, a is the size of the receptive field center, b is the size of

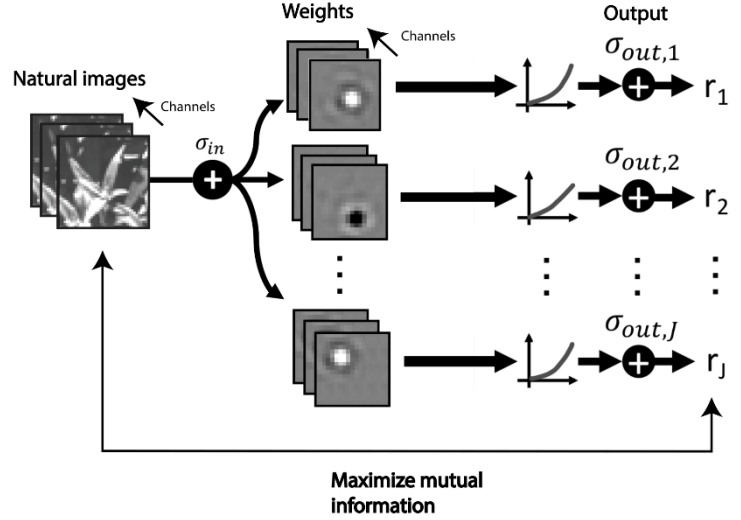


Figure 3: Efficient coding model architecture. Natural images (plus input noise σ_{in}) are multiplied by linear filters across multiple channels. These channels represent either cone inputs (Aim 1) or latencies (Aim 2). The output is perturbed by noise σ_{out} and rectified by a non-linear function, resulting in firing rate r_j for model neuron j . The weights of the filters are optimized to maximize the mutual information between the natural images and the firing rates across all neurons.

the receptive field surround. d and c are the relative strength of the center and the surround for each color channel, respectively. To allow these values to converge, d is restrained to have an L2-norm of 1. a and b are both restricted to have positive values, and c is restricted to have values ranging from 0 to 1. On top of facilitating training, this parametrization also allows us to directly characterize how each model RGCs integrate L, M and S inputs.

I will model the fovea versus the periphery by changing the number cone-RGC ratio of the model. The fovea will be modeled by having 3 times more RGCs than cones, while the periphery will be modeled by having the same number of cones and RGCs.

Expected outcomes

I expect model RGCs to have three different receptive field types that are analogous to parasol, midget and bistratified cells. Based on previous research^{12, 13, 15}, each cell type should form a mosaic that tiles the entire visual field. We can categorize which RGC type a model neuron by analyzing the d parameter for L, M and S inputs. Similar to parasol cells, I expect some neurons will have L, M and S inputs that are all positive or all negative, respectively. I expect another group of neurons to have positive S cones and negative L and M cones, similarly to bistratified cells. Finally, there should be another group of neurons that looks similar to midget cells. With a high RGC-cones ratio (>3), these cells should have opponency between L and M cones, and sum L and M cones with lower RGC-cones ratios (~ 1).

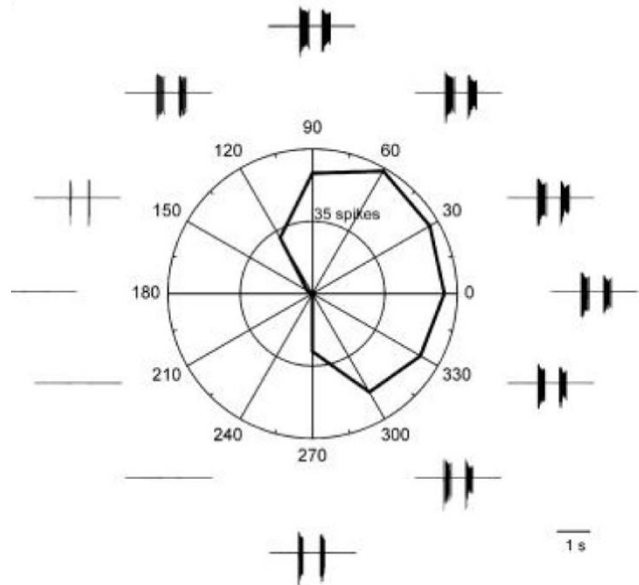


Figure 4: Responses of DSGC to a rectangle moving in 12 directions; both the polar plot and spike traces demonstrate strong directional responses. Taken from Weng, Sun & He (2005)

Aim 2: Expand efficient coding models to explain motion-selectivity in RGCs

Background and rationale

Both the world and ourselves are constantly in movement, which makes motion an important aspect of visual scenes. The retina encodes motion by modulating its response over time: Stimuli that initially increases the firing rate of a neuron typically will typically decrease it at later latencies. How fast this modulation occurs – the neuron's temporal frequency – changes across different neuron types. Midget cells encode low temporal frequencies while parasol cells encode high temporal frequencies, which allows these two populations to encode objects moving at different angular velocities on our visual fields. The retina also needs to encode the direction an object is moving in, which is done by Direction-Selective Ganglion Cells (DSGC). DSGC are divided into four subtypes that encode each of the four cardinal directions (up, down, left, right)²⁹. Efficient coding models have successfully explained why midget and parasol cells encode different spatial and temporal frequencies. However, these models had constraints that prevented them from reproducing direction-selectivity in RGCs¹². In Aim 2 I will build an efficient coding model that reproduces direction-selectivity. The contribution of this project will be a theoretical understanding of retinal direction selectivity from efficient coding principles.

Experimental method and design

I will build an efficient coding model to encode natural movies from the Chicago Motion Database²⁷. This model will be conceptually similar to the one from Aim 1 (Figure X). The main difference

is in the input channels, which will be different latencies instead of different colors. At time t , the model neurons will receive inputs from the image at time t , $t - 1$, $t - 2$, etc. This reflects how RGCs don't instantaneously fire in response to stimuli but instead integrate their inputs over time³². The weights from each latency will be estimated separately to maximize the mutual information between the inputs and outputs of the model (see Aim 1). The weights will be parametrized by difference-of-gaussians, similarly to Aim 1. We will measure the amount of direction-selectivity in each neuron by computing how much the kernel centers change across latencies.

Expected outcomes

I expect that this new efficient coding model will have a subset of neurons that are direction-selective. These neurons should be divided into four different types, with each type encoding one of the four cardinal directions. Each type should have neurons that form a mosaic and tile the entire visual space.

References

1. Atick, J.J., Z. Li, and A.N. Redlich, *Understanding retinal color coding from first principles*. Neural computation, 1992. **4**(4): p. 559-572.
2. Atick, J.J. and A.N. Redlich, *Towards a theory of early visual processing*. Neural computation, 1990. **2**(3): p. 308-320.
3. Atick, J.J. and A.N. Redlich, *What does the retina know about natural scenes?* Neural computation, 1992. **4**(2): p. 196-210.
4. Balasubramanian, V. and M.J. Berry II, *A test of metabolically efficient coding in the retina*. Network: Computation in Neural Systems, 2002. **13**(4): p. 531.
5. Barlow, H.B., *Possible principles underlying the transformation of sensory messages*. Sensory communication, 1961. **1**(01): p. 217-233.
6. Baylor, D., B. Nunn, and J. Schnapf, *Spectral sensitivity of cones of the monkey *Macaca fascicularis**. The Journal of Physiology, 1987. **390**(1): p. 145-160.
7. Conway, B.R., et al., *Advances in color science: from retina to behavior*. Journal of Neuroscience, 2010. **30**(45): p. 14955-14963.
8. Crook, J.D., et al., *Horizontal cell feedback without cone type-selective inhibition mediates "red-green" color opponency in midget ganglion cells of the primate retina*. Journal of Neuroscience, 2011. **31**(5): p. 1762-1772.
9. Dacey, D.M. *Physiology, morphology and spatial densities of identified ganglion cell types in primate retina*. in *Ciba Foundation Symposium 184-Higher-Order Processing in the Visual System: Higher-Order Processing in the Visual System: Ciba Foundation Symposium 184*. 2007. Wiley Online Library.
10. Doi, E., et al., *Efficient coding of spatial information in the primate retina*. Journal of Neuroscience, 2012. **32**(46): p. 16256-16264.
11. Doi, E., et al., *Spatiochromatic receptive field properties derived from information-theoretic analyses of cone mosaic responses to natural scenes*. Neural computation, 2003. **15**(2): p. 397-417.
12. Jun, N.Y., G. Field, and J. Pearson, *Efficient coding, channel capacity, and the emergence of retinal mosaics*. Advances in neural information processing systems, 2022. **35**: p. 32311-32324.
13. Jun, N.Y., G.D. Field, and J. Pearson, *Scene statistics and noise determine the relative arrangement of receptive field mosaics*. Proceedings of the National Academy of Sciences, 2021. **118**(39): p. e2105115118.
14. Kaneko, A., *Physiology of the retina*. Annual review of neuroscience, 1979. **2**(1): p. 169-191.

15. Karklin, Y. and E. Simoncelli, *Efficient coding of natural images with a population of noisy linear-nonlinear neurons*. Advances in neural information processing systems, 2011. **24**.
16. Kim, J.S., et al., *Space-time wiring specificity supports direction selectivity in the retina*. Nature, 2014. **509**(7500): p. 331-336.
17. Kingma, D.P. and J. Ba, *Adam: A method for stochastic optimization*. arXiv preprint arXiv:1412.6980, 2014.
18. Lee, T.-W., T. Wachtler, and T.J. Sejnowski, *Color opponency is an efficient representation of spectral properties in natural scenes*. Vision Research, 2002. **42**(17): p. 2095-2103.
19. Levenstein, D., et al., *On the role of theory and modeling in neuroscience*. Journal of Neuroscience, 2023. **43**(7): p. 1074-1088.
20. Lewicki, M.S., *Efficient coding of natural sounds*. Nature neuroscience, 2002. **5**(4): p. 356-363.
21. Maheswaranathan, N., et al., *Inferring hidden structure in multilayered neural circuits*. PLoS computational biology, 2018. **14**(8): p. e1006291.
22. Martin, P.R., et al., *Chromatic sensitivity of ganglion cells in the peripheral primate retina*. Nature, 2001. **410**(6831): p. 933-936.
23. Miller, L.E., et al., *Somatosensory cortex efficiently processes touch located beyond the body*. Current Biology, 2019. **29**(24): p. 4276-4283. e5.
24. Ocko, S., et al., *The emergence of multiple retinal cell types through efficient coding of natural movies*. Advances in Neural Information Processing Systems, 2018. **31**.
25. Pitkow, X. and M. Meister, *Decorrelation and efficient coding by retinal ganglion cells*. Nature neuroscience, 2012. **15**(4): p. 628-635.
26. Rasmussen, R. and K. Yonehara, *Contributions of retinal direction selectivity to central visual processing*. Current Biology, 2020. **30**(15): p. R897-R903.
27. Salisbury, J.M. and S.E. Palmer, *Optimal prediction in the retina and natural motion statistics*. Journal of Statistical Physics, 2016. **162**(5): p. 1309-1323.
28. Soto, F., et al., *Efficient coding by midget and parasol ganglion cells in the human retina*. Neuron, 2020. **107**(4): p. 656-666. e5.
29. Vaney, D.I., B. Sivy, and W.R. Taylor, *Direction selectivity in the retina: symmetry and asymmetry in structure and function*. Nature Reviews Neuroscience, 2012. **13**(3): p. 194-208.
30. Wässle, H., et al., *Retinal ganglion cell density and cortical magnification factor in the primate*. Vision research, 1990. **30**(11): p. 1897-1911.
31. Wei, W., et al., *Development of asymmetric inhibition underlying direction selectivity in the retina*. Nature, 2011. **469**(7330): p. 402-406.
32. Yasui, S., W. Davis, and K.-I. Naka, *Spatio-temporal receptive field measurement of retinal neurons by random pattern stimulation and cross correlation*. IEEE Transactions on Biomedical Engineering, 1979(5): p. 263-272.