<u>Introduction to Revisions</u>

I would like to thank my reviewers for their helpful and constructive feedback. I have made many changes that address my reviewer's concerns, which in turn also strengthens my proposal. The constructive critiques have prompted me to add many details about my project, which should make the proposal clearer. I have outlined the primary concerns below and explain how I addressed them.

"Lack of alternative hypotheses or how unexpected outcomes will be interpreted":
I have added an alternative hypotheses section for both Aim 1 and Aim 2. In Aim 1, I explain how we will use a mathematically tractable version of our model if we fail to find color opponency. This model would be relatable to previous literature and allow us to understand the parameter requirements to obtain color opponency. In Aim 2, I explain how we can train our model on different samples of natural movies to understand what natural stimuli makes the encoding direction-selectivity efficient.

"Lack of cohesion between Aim 1 and Aim 2":
Aim 1 deals input from different chromatic channels, and Aim 2 deals with inputs from different latencies. While this was previously unclear, what these Aims have in common is that they involve multiple correlated channels as inputs. Previous efficient coding models have mostly focused on a single channel, and it is unclear how efficient coding strategies change when dealing with strongly correlated input channels. I have outlined this theme more explicitly in the Specific Aims. The efficient coding models for Aim 1 and Aim 2 are very similar, since they deal with similar problems. For this reason, the description of the efficient coding model is now its own section instead of a part of Aim 1.

"The literature review of Direction-Selective Ganglion Cells in Aim 2 is weak":
I now give a more thorough literature review of Direction-Selective Ganglion Cells (DSGCs) in Aim 2. I distinguish ON-OFF DSGCs from ON DSGCs, and explain how their receptive fields properties differ from each other. The expected outcomes for this Aim have also been extended to explain what experimental findings we aim to replicate.

"Concerns about the parametrization of filters as Difference-Of-Gaussians":
The reviewers were concerned by the Difference-Of-Gaussian parametrization of the filters. For both Aim 1 and Aim 2, preliminary analyses have shown that the filters consistently converge to Difference-Of-Gaussians. We therefore believe that this parametrization will not affect the model's results. In light of the reviewers' concerns, we will perform follow-up experiments after the completion of each Aim to verify that the results can be generalized to models with filters that are not parametrized.

"The efficient coding model could be explained better":
I've included a general introduction to efficient coding in the approach section. This includes the definition of mutual information and entropy, which are essential aspects of efficient coding. I've also given a more mathematically rigorous explanation of the model, where I describe in more details how noise is included and what the output nonlinearity is.

"Spatial downsampling of natural images may lose important statistical relationships between locations":
How we draw training samples from the full images was previously unclear. Instead of downsampling, we crop smaller 16x16 training samples from the full 1000x1280 images. This allows us to create a large number of training samples from each image. We've now clarified this section of the approach.

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 process. 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[1-6], audition[7] and touch[8]. 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[1, 2, 6, 9, 10]. However, we still lack efficient coding predictions for how the retina processes strongly correlated inputs across multiple channels. Solving this problem is important to understand 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, respectively[11]. 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 cells[12], which, in the fovea, are 'red-green opponent' and encode a contrast between L and M inputs[11-13]. In the periphery, these midget cells are not red-green opponent but instead sum L and M cone inputs[12-14]. 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 cone[15]. 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[16, 17]. The major type of motion encoding RGCs are ON-OFF direction-selective ganglion cells[18], which are divided into four subtypes that respond preferentially to each of the four cardinal directions[17]. 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 fields[19], 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 directions[17]. 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) explanations[20]. The retina is fairly well-understood both from descriptive and mechanistic perspectives[12, 21]. 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[12, 21]: 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 sends 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.
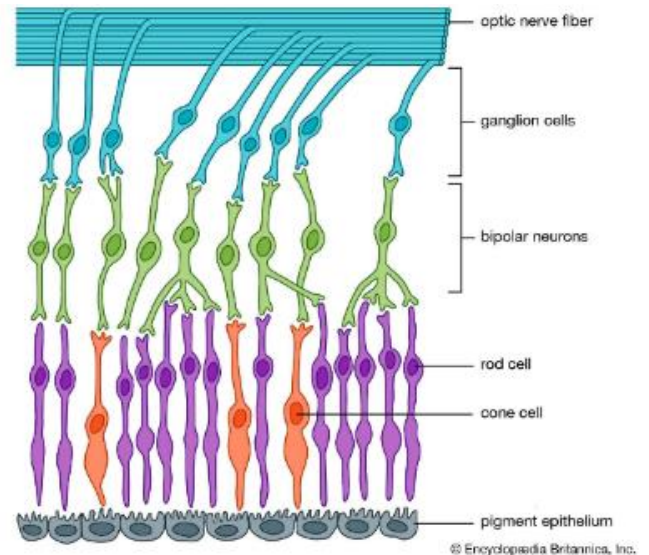

Figure 1: Schematic structure of the retina

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[9, 10]. 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[3, 6, 19, 22, 23], 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. linearity[24] and space-time separability[19]) which are difficult to relate to retinal physiology[13, 25].

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. 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.

However, such models have previously been either limited (1) to one single input channel[6, 22] or (2) by making strong assumptions about how receptive fields should change across channels[19]. The goal of this proposed project is to extend such models to natural stimuli that have multiple correlated channels. , 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).

Efficient coding model

The efficient coding model we use will be similar to other work from our lab[19, 22]. The inputs will consist of D pixels patches of Natural Images X corrupted by gaussian input noise $n_{in} \sim N(0, C_{in})$. The model RGC j takes as input a linear combination of these natural images:

$$y_j = w_j^T (x_j + n_{in})$$

Where w_j are unit-norm kernels ($\|w_j\| = 1$) and represent how much each photoreceptor contributes to the response of model RGC j. y_j is then passed through a softplus nonlinearity:

$$\eta(y_j) = \frac{log(1 + e^{\beta y_j})}{\beta}$$

And further corrupted by output noise $n_{out} \sim N(0, C_{out})$ to produce firing rate r_j with threshold θ_j and gain γ_j:

$$r_j = \gamma_j \cdot \eta(y_j - \theta_j) + n_{out}$$

The above parameters are optimized via Adam to maximize the mutual information between the inputs X and the outputs R, under a firing rate constraint. Mutual information represents the amount of information, in bits, that is transmitted to RGCs from photoreceptors. Mutual information is equivalent to subtracting the entropy of the natural images X from the conditional entropy of X on R:

$$I(X; R) \equiv H(X) - H(X|R)$$

Where the entropy H(X) is defined as:

$$H(X) = - \sum_{x \in X} p(x) log(p(x)) = \mathbb{E}[-log(p(X))]$$

Previous work has derived a closed-form solution for the mutual information between X and R[6, 19, 22]. The above parameters are optimized to maximize this mutual information:

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

$$\text{subject to } \mathbb{E}[r_j] = 1$$

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 W is the weight matrix. 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 softmax function, the diagonal of G has binary values (1 if the neuron is firing and 0 if not). This function is maximized using Adam optimization[26]. To represent the metabolic cost of firing spikes, each neuron is be restricted to have a fixed average firing rate.

In both aims, we will train the model on multiple correlated input channels (color in Aim 1, and time in Aim 2). Having multiple input channels drastically increases the number of parameters, which can make the model significantly harder to optimize. However, preliminary results have shown that these filters consistently converge to difference-of-gaussians even when trained to encode different channels that represent either color or time. We can use this fact to our advantage to facilitate training by reducing the number of parameters. To that effect, we will parametrize the linear filters of each neuron to be difference-of-gaussians:

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

Where a, b, c, d and z are vectors of length k that represent each of the k different channels. z is the distance from the receptive field center, a is the size of the receptive field center, b is the size of 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 either color or time across different channels.

Aim 1: Expand efficient coding models to encompass chromatic information

<u>Background and rationale</u>

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
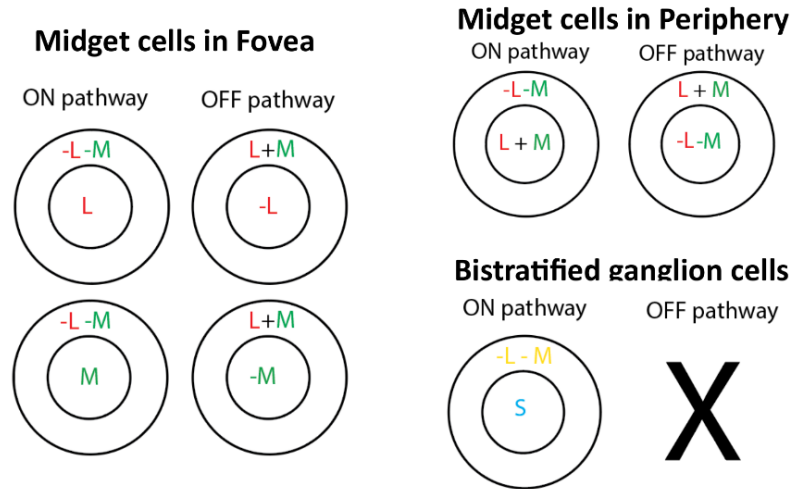

Figure 2: Receptive fields of midget and bistratified ganglion cells.

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 three[24]. Independent-component analysis can explain why color-opponency with three cones is efficient[27], 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 differently by midget, parasol and bistratified cells. Aim 1 will address this question by building efficient coding models that encode chromatic natural images.
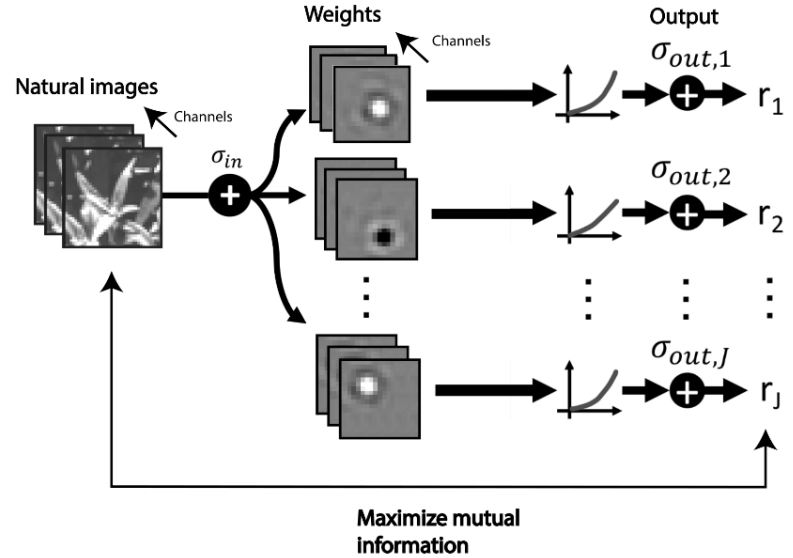

Figure 3: Efficient coding model architecture. Natural images (plus input noise σ_{in}) are multipled by linear filters across multiple channels. These channels represent either cone inputs (Aim 1) or latencies (Aim 2). The output is perturbed by noise σ_{out} and rectified by a non-linear function, 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.

Experimental methods and design

To study how to efficiently encode chromatic natural images, we will use the Kyoto Natural Images Dataset[28]. 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 sensitivities[28]. These responses were further transformed with an empirical cone non-linear function[29]. Each of these images have 1000x1280x3 pixels. We will crop these images into smaller training samples of 16x16x3 pixels.

We will train the efficient coding model described above on the Kyoto Natural Images Dataset. We will analyze the learned parameters of the model, specifically the d parameter, to understand how neurons cluster into different types that process chromatic information differently. We will change the RGC-cones ratio to test whether this ratio changes how midget cells process chromatic information, and compare these results to known receptive fields in the fovea and periphery.

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[6, 19, 22], 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).

Alternative approaches and hypotheses

Due to the high correlations across color channels, it is possible that receptive fields converge on adding different colors channels instead of subtracting them. If that occurs, we will build a linear version of our model that is mathematically tractable. We will then compare our results to previous literature that found color-opponency to be efficient by making assumptions about the Fourier structure of chromatic natural images[24]. This mathematically tractable model should give us an intuition about what conditions are sufficient and necessary for different chromatic receptive fields to be efficient.

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

Background and rationale

Both the world and ourselves are constantly in movement, which makes motion an important aspect of visual scenes. Consistent with this idea, RGCs respond strongly to moving stimuli. These responses vary as a function of the
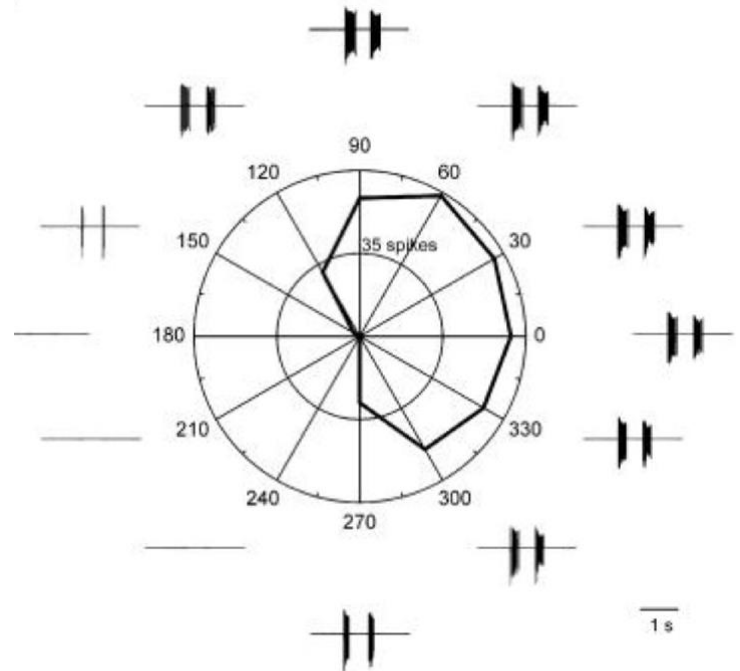

Figure 4: Responses of DSGC to a rectangle moving in 12 directions; both the polar plot and spike traces demonstrate strong directional responses. Taken from Weng, Sun & He (2005)

orientation of motion, and the optimal orientation is usually either pointing towards the fovea or perpendicular to it. This orientation-bias is present in both midget and parasol cells[30]. However, these neurons can't individually discriminate the direction of motion: Their responses would be similar if an object is moving left-to-right than right-to-left. Only a subset (~12%) of neurons are direction-selective[17]. These Direction-Selective Ganglion Cells (DSGCs) are divided into two main types: ON-OFF DSGCs and ON DSGCs. ON-OFF DSGCs are divided into four subtypes that encode each of the four cardinal directions (anterior, posterior, superior and inferior)[17]. ON-OFF DSGCs respond to the onset of both ON and OFF stimuli, a property similar to complex cells in the primary visual cortex[31]. These cells were initially discovered in the rabbit retina[32], but have also more recently been discovered in the mouse[17] and macaque retina[33]. While ON-OFF DSGCs have relatively small receptive fields and detect local motion, ON DSGCs have larger receptive fields that encode global motion. ON DSGCs are divided into three subtypes that encode the anterior, posterior and inferior directions of motion[17]. ON DSGCs respond to light (but not dark stimuli), and are divided into two different populations that have either transient or sustained responses[17]. ON DSGCs are also less numerous than their ON-OFF counterparts. These cells have been discovered in the rabbit and mouse, but have yet to be found in macaques.

Previous work from our lab built efficient coding models for spatiotemporal receptive fields. However, these models had constraints that prevented them from having direction-selective neurons[19]. In Aim 2 I will test the hypothesis that having a subset of direction-selective cells is efficient, due to the predominance of motion in natural movies. To test this hypothesis, I need an efficient coding model that can integrate multiple correlated channels – each channel representing a different latency. The model previously described is a great candidate to solve this problem, as it makes few assumptions about the temporal structure of receptive fields. 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 Database[34]. This model will be conceptually similar to the one from Aim 1 (Figure 3). 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 time[35]. 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.

To encapsulate the responses properties of ON-OFF DSGCs, we will change the output nonlinearity from a softmax function to a Parametric Rectified Linear Unit:

$$\eta(y_j) = max(0, y_j) + a_j min(0, y_j)$$

This output non-linearity is conceptually similar to the energy model for complex cells in the primary visual cortex[31].

Expected outcomes

Previous work from my lab[19, 22] suggests that new RGCs types emerge as we increase the number of neurons. Most of RGCs in the retina are not DSGCs, which is why I expect that the first neurons to emerge will not be DSGCs. Instead, these should have similar spatiotemporal properties to midget and parasol cells[19]. However, as we increase the number of neurons, I expect that we will eventually find ON-OFF DSGCs, followed by ON DSGCs. ON-OFF DSGCs should have relatively small receptive fields and encode the four cardinal directions. Moreso, the a_j parameter from the output nonlinearity should be negative, allowing these neurons to respond to both ON and OFF stimuli. ON DSGCs should have larger receptive fields, and encode motion in the anterior, posterior and inferior direction. The estimated a_j should be near zero, allowing these neurons to respond to ON (but not OFF) onset.

The known physiology about DSGCs is fairly detailed. It is possible we will not be able to reproduce all of those experimental findings from efficient coding principles. However, we hope that we will be able to show that ON-OFF DSGCs in the four cardinal directions can be derived from efficient coding principles. Other than that, it will be of interest to see which findings we can replicate from efficient coding principles, and which ones we cannot.

Alternative approaches and hypotheses

Because DSGCs are only a small fraction (~12%) of RGCs, it might require a large RGC-cones ratio to obtain neurons that are directions-selective. If we fail to replicate DSGCs with the computing resources we have, we will remove stationary sections of the movies from our dataset, and train the model exclusively on moving stimuli. We will also try reducing the number of learned parameters; for example, by assuming a_j is negative or assuming a, b, c and d do not vary across channels.

References

1. 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.
2. Doi, E., et al., *Efficient coding of spatial information in the primate retina.* Journal of Neuroscience, 2012. **32**(46): p. 16256-16264.
3. 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**.
4. Pitkow, X. and M. Meister, *Decorrelation and efficient coding by retinal ganglion cells.* Nature neuroscience, 2012. **15**(4): p. 628-635.
5. Soto, F., et al., *Efficient coding by midget and parasol ganglion cells in the human retina.* Neuron, 2020. **107**(4): p. 656-666. e5.
6. 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**.
7. Lewicki, M.S., *Efficient coding of natural sounds.* Nature neuroscience, 2002. **5**(4): p. 356-363.
8. Miller, L.E., et al., *Somatosensory cortex efficiently processes touch located beyond the body.* Current Biology, 2019. **29**(24): p. 4276-4283. e5.
9. Atick, J.J. and A.N. Redlich, *Towards a theory of early visual processing.* Neural computation, 1990. **2**(3): p. 308-320.
10. Atick, J.J. and A.N. Redlich, *What does the retina know about natural scenes?* Neural computation, 1992. **4**(2): p. 196-210.
11. 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.
12. 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.
13. Conway, B.R., et al., *Advances in color science: from retina to behavior.* Journal of Neuroscience, 2010. **30**(45): p. 14955-14963.
14. Martin, P.R., et al., *Chromatic sensitivity of ganglion cells in the peripheral primate retina.* Nature, 2001. **410**(6831): p. 933-936.
15. Wässle, H., et al., *Retinal ganglion cell density and cortical magnification factor in the primate.* Vision research, 1990. **30**(11): p. 1897-1911.
16. Rasmussen, R. and K. Yonehara, *Contributions of retinal direction selectivity to central visual processing.* Current Biology, 2020. **30**(15): p. R897-R903.
17. Vaney, D.I., B. Sivyer, 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.
18. Wei, W., et al., *Development of asymmetric inhibition underlying direction selectivity in the retina.* Nature, 2011. **469**(7330): p. 402-406.
19. 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.
20. Levenstein, D., et al., *On the role of theory and modeling in neuroscience.* Journal of Neuroscience, 2023. **43**(7): p. 1074-1088.
21. Kaneko, A., *Physiology of the retina.* Annual review of neuroscience, 1979. **2**(1): p. 169-191.
22. 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.
23. Maheswaranathan, N., et al., *Inferring hidden structure in multilayered neural circuits.* PLoS computational biology, 2018. **14**(8): p. e1006291.

24. Atick, J.J., Z. Li, and A.N. Redlich, *Understanding retinal color coding from first principles.* Neural computation, 1992. **4**(4): p. 559-572.
25. Kim, J.S., et al., *Space–time wiring specificity supports direction selectivity in the retina.* Nature, 2014. **509**(7500): p. 331-336.
26. Kingma, D.P. and J. Ba, *Adam: A method for stochastic optimization.* arXiv preprint arXiv:1412.6980, 2014.
27. 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.
28. 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.
29. 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.
30. Antinucci, P. and R. Hindges, *Orientation-selective retinal circuits in vertebrates.* Frontiers in neural circuits, 2018. **12**: p. 11.
31. Mechler, F. and D.L. Ringach, *On the classification of simple and complex cells.* Vision research, 2002. **42**(8): p. 1017-1033.
32. Barlow, H. and W.R. Levick, *The mechanism of directionally selective units in rabbit's retina.* The Journal of physiology, 1965. **178**(3): p. 477.
33. Kim, Y.J., et al., *Origins of direction selectivity in the primate retina.* Nature communications, 2022. **13**(1): p. 2862.
34. 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.
35. 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.