

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<sup>1</sup>. 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<sup>2-7</sup>, audition<sup>8</sup> and touch<sup>9</sup>. 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<sup>2, 3, 7, 10, 11</sup>. 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.

Retinal processing of visual information follows a well-known structure<sup>12, 13</sup>: 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. and send 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.

Early theoretical work on efficient coding explained how this center-surround organization arises from decorrelation, both for achromatic and for color inputs<sup>10, 11</sup>, but 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<sup>4, 7, 14-16</sup>, 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<sup>17</sup> and space-time separability<sup>15</sup>) which are difficult to relate to retinal physiology<sup>18, 19</sup>. To close this gap, I will use more general efficient coding models to model how RGCs process color information (Aim 1) and motion (Aim 2). In doing so, I will take advantage of the established collaboration between my supervisor and Dr. Greg Field (UCLA) on efficient coding models of the retina<sup>14, 15</sup>. This collaboration also opens the possibility of testing the predictions of our models against new experimental data.

### **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<sup>20</sup>. 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<sup>13</sup>, which, in the fovea, are ‘red-green opponent’ and encode a contrast between L and M inputs<sup>13, 19, 20</sup>. In the periphery, these midget cells are not red-green opponent but instead sum L and M cones inputs<sup>13, 19, 21</sup>. Why midget cells use different coding strategies for the

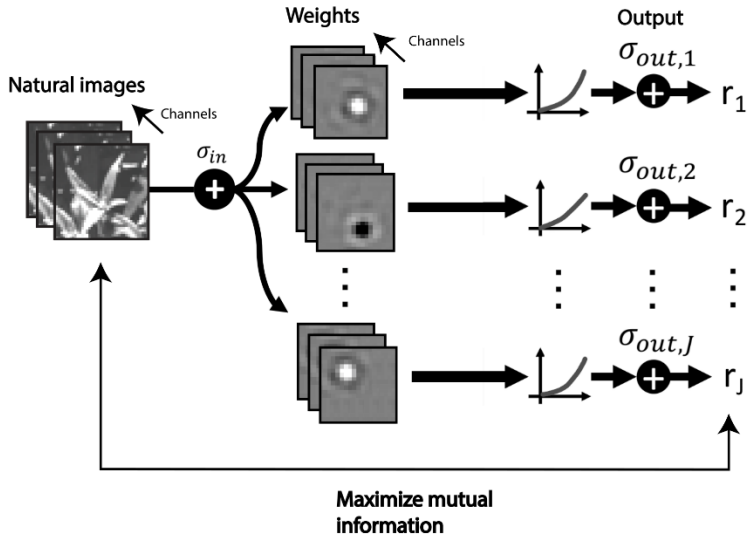


Figure 1: Efficient coding model architecture. Natural images (plus input noise  $\sigma_{in}$ ) are multiplied by linear filters across multiple channels. These channels represent either cone inputs (Aim 1) or latencies (Aim 2). The output is perturbed by noise  $\sigma_{out}$  and rectified by a non-linear 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 total firing rate across all neurons<sup>7, 10, 14, 15</sup>. To model differences between the fovea and the periphery, we will change the number of outputs RGCs relative to the number of input cones. Preliminary results suggest that if we fix the RGC-cone ratio at 1:1, efficient coding predicts that L and M inputs should be summed. We will next test if we can replicate chromatic receptive fields in the fovea by building efficient coding models with a 3:1 RGC-cone ratio. 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 allow us to assess whether efficient coding can accurately explain how the retina encodes chromatic information.

### Aim 2: Expand efficient coding models to explain why some RGCs are motion-selective

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<sup>23, 24</sup>. The major type of motion encoding RGCs are ON-OFF direction-selective ganglion cells<sup>25</sup>, which are divided into four subtypes that respond preferentially to each of the four cardinal directions<sup>24</sup>. 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<sup>15</sup>, 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. This model will be conceptually similar to the one from Aim 1, where the inputs to model RGCs come from multiple channels, each of which representing a different latency. 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<sup>24</sup>. 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.

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<sup>22</sup>. Here we will test that hypothesis by building an efficient coding model for chromatic natural images (Figure 1). The inputs to this model will be the LMS cone responses to a series of natural images. We will then filter these responses through three different spatial filters (one for each cone), which are summed and then followed with an output non-linearity. The weights of the filter will be fit to optimize mutual information between the cone inputs and the RGC outputs. Consistent with efficient coding principles, there will be a constraint on