Do we need space and time to understand color vision?

David St-Amand[1], Greg Field[2] and John Pearson[1,3]

1. Duke Department of Neurobiology 2. UCLA Department of Neurobiology
3. Duke Department of Electrical and Computer Engineering

Introduction

Efficient coding models have been especially successful at explaining how populations of neurons should encode achromatic natural images (Karklin & Simoncelli, 2011; Jun et al., 2021), but efficient coding predictions about encoding color are still unclear.

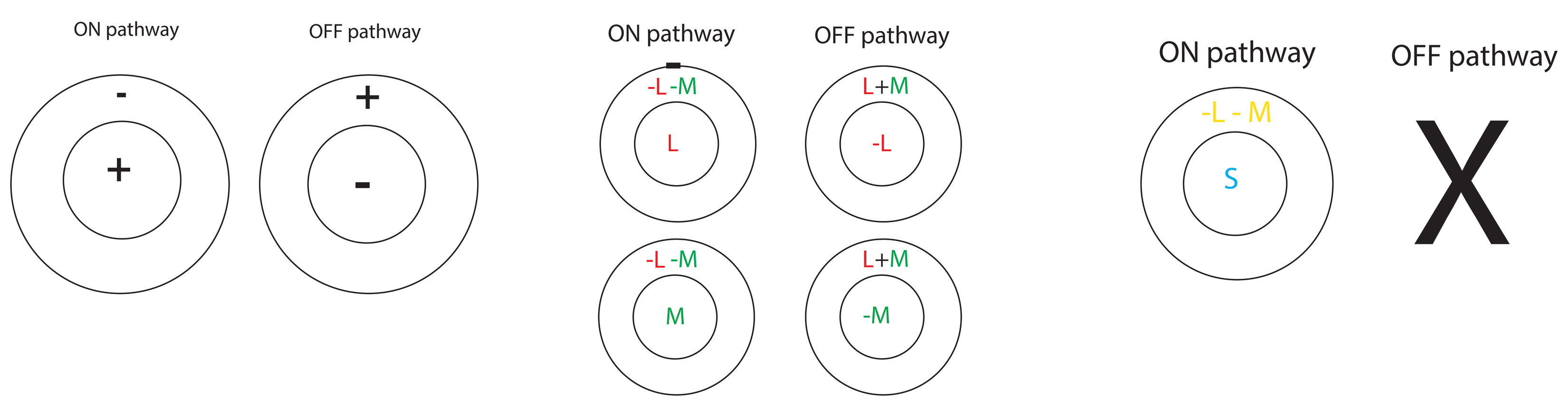
Retinal ganglion cells (RGCs) are divided into three main classes that encode separate chromatic information about natural scenes. The receptive fields of each class not only have different chromatic properties, but are also tuned to different spatial and temporal frequencies.

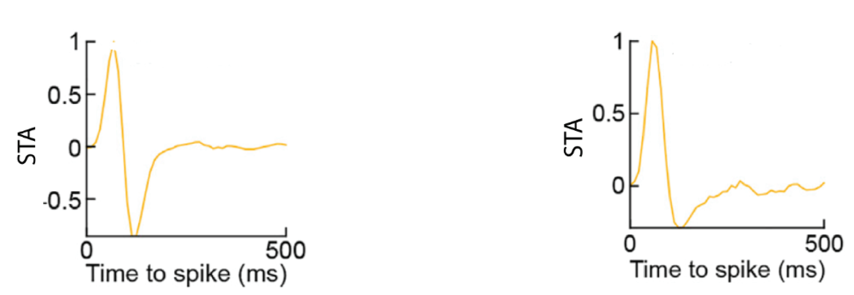
Previous spatiotemporal efficient coding models have successfully explained the segregation of RGCs into different types, by suggesting a tradeoff between spatial and temporal frequencies. Here we hypothesize that such models can also explain why – and how- different cell types encode chromatic information.

Parasol cells (5-20%):
1. Black/white opponent
2. Low spatial frequency
3. High temporal frequency

Midget cells (50-90%):
1.Red/green opponent
2. High spatial frequency
3. Low temporal frequency

Bistratified cells (1-10%):
1.Blue/Yellow opponent
2. Low spatial frequency (?)
3. Low temporal frequency (?)





Methods

1. Convert natural movies (e.g. documentary) from RGB to LMS coordinates assuming D65 luminance and using CIE standards.
2. Compute the Fourier Transform of natural movies and the covariance matrix for each channel pair:

$$F_c(k_x, k_y, \omega) = \iint_{-\infty}^{+\infty} f(x, y, t, c) e^{i(xk_x + yk_y + t\omega)} dxdydt \qquad C_x(k, \omega, c_1, c_2) = F_{c_1}(k, \omega) F_{c_2}(k, \omega)^*$$

k: *Spatial frequency* $\quad$ ω: *Temporal frequency* $\quad$ c: *Color channel* $\quad$ $k = \sqrt{k_x^2 + k_y^2}$

3. Perform eigendecomposition on Cx and use its eigenvalues to find the filters that optimize mutual information between inputs and ouptuts:

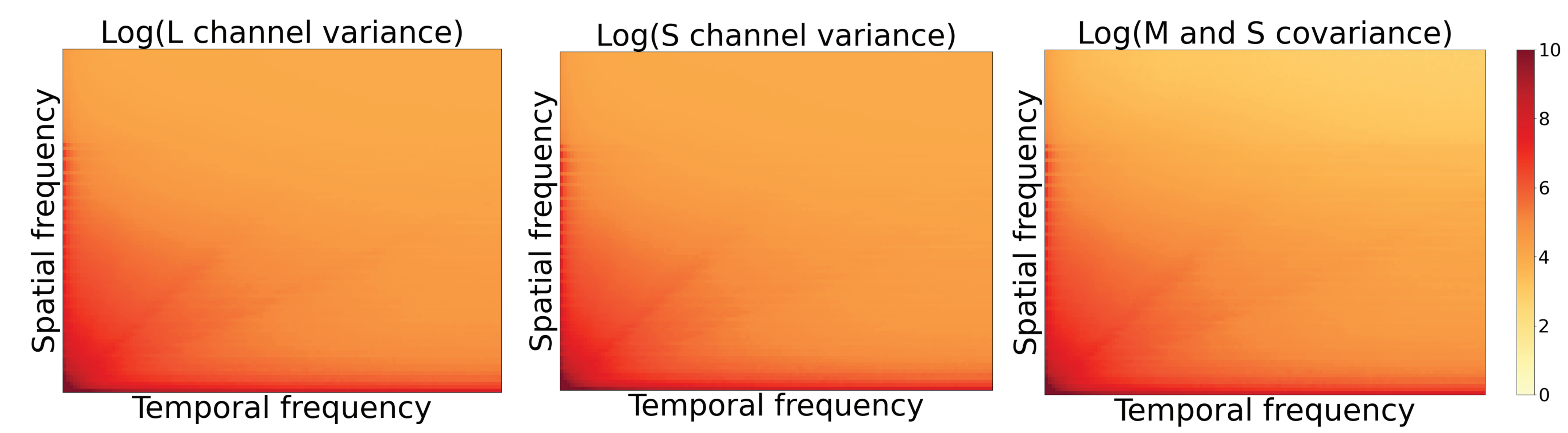
$$|a_p(k,\omega)|^2 = \sigma_{\text{out}}^2 \left[\frac{1}{2} \frac{\lambda_p(k,\omega)}{\lambda_p(k,\omega)+1} \left(\sqrt{1 + \frac{4}{\sigma_{\text{out}}^2 \nu_p \lambda_p(k,\omega)}} + 1 \right) - 1 \right]_+$$

$\lambda_c(k,\omega)$: c^{th} *eigenvalue of* $C_x(k,\omega)$ $\quad$ $|a_c(k,\omega)|^2$: *Optimal filter power for eigenchannel c*

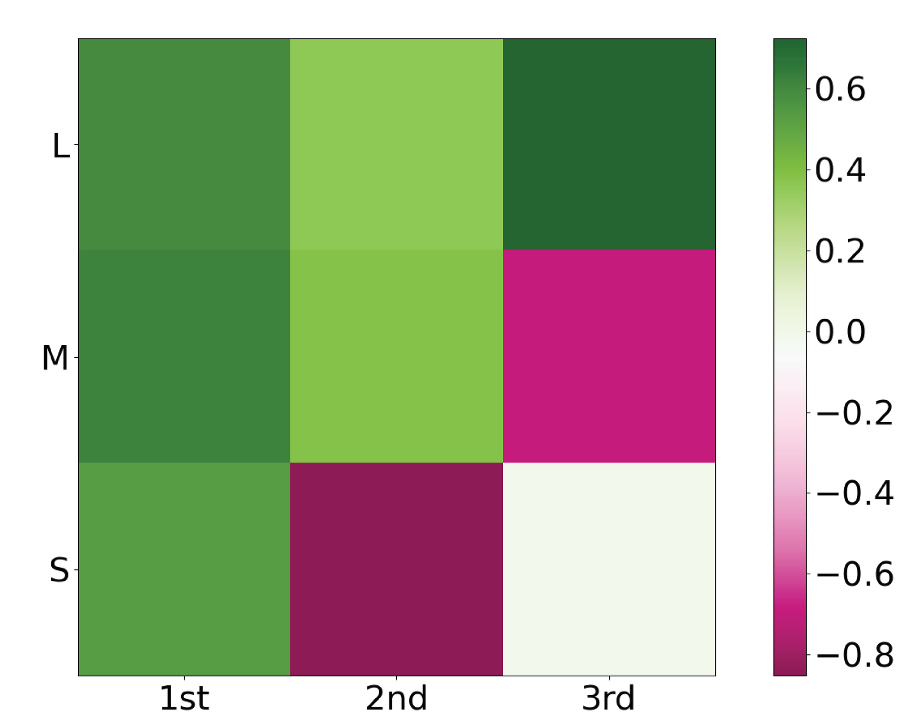
σ_{out}^2: *Output noise*

Results

Cx has similar distributions across channels

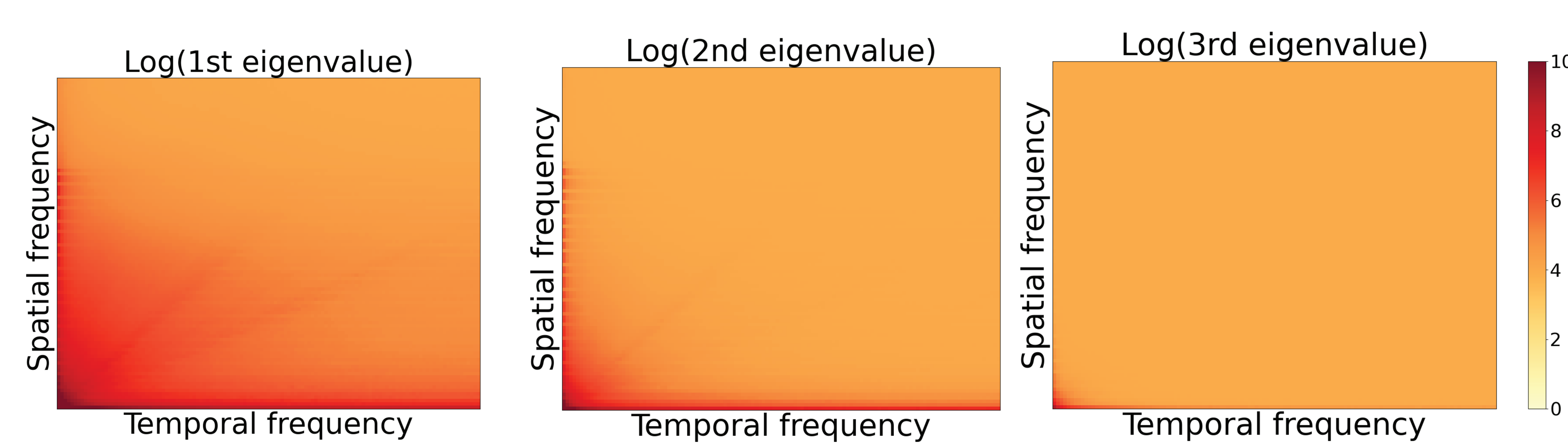


Eigendecomposition naturally recovers the chromatic structure of parasol, bistratified and midget cells



First eigenvector: L + M + S
Second eigenvector: L + M – S
Third eigenvector: L – M

The second and third eigenvalues only have power at low spatial and low temporal frequencies



Discussion

1. We found that luminance, blue/yellow and red/green opponency have different spatiotemporal distributions across natural movies.
2. The next step is to use these eigenvalues to infer the optimal filters that maximize mutual information, and compare these to the retinal biology.
3. Because the optimal strategy to optimize mutual information is to whiten the inputs, we expect the model to predict a larger number of red/green opponent RGCs that are selective to high spatial frequencies.