Data Visualization

2017-01-24

Review From Yesterday

- Purposes of visualization
- Attribute domain taxonomy
- Graphical Integrity
- Maximizing impact

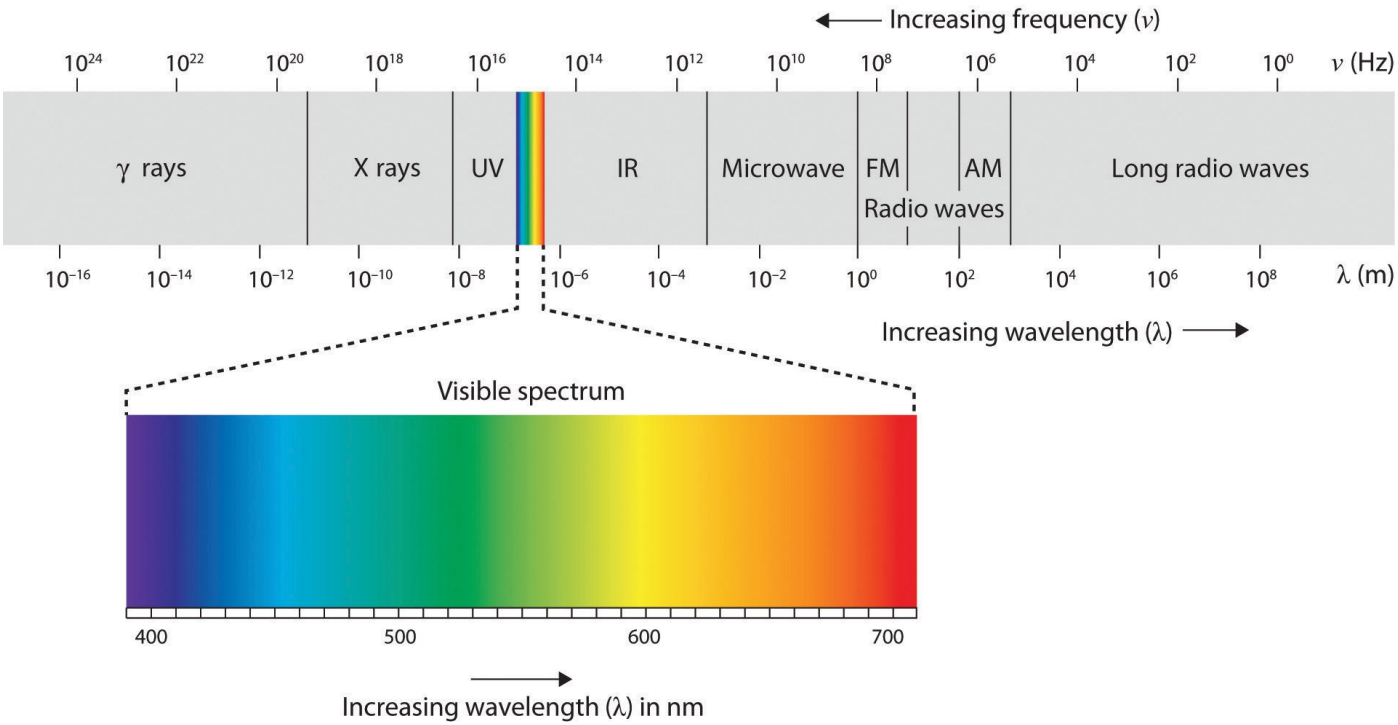
Today

- Go over assignment
- Encoding data as color
- Visualizing distributions
- Exploring high dimensional data

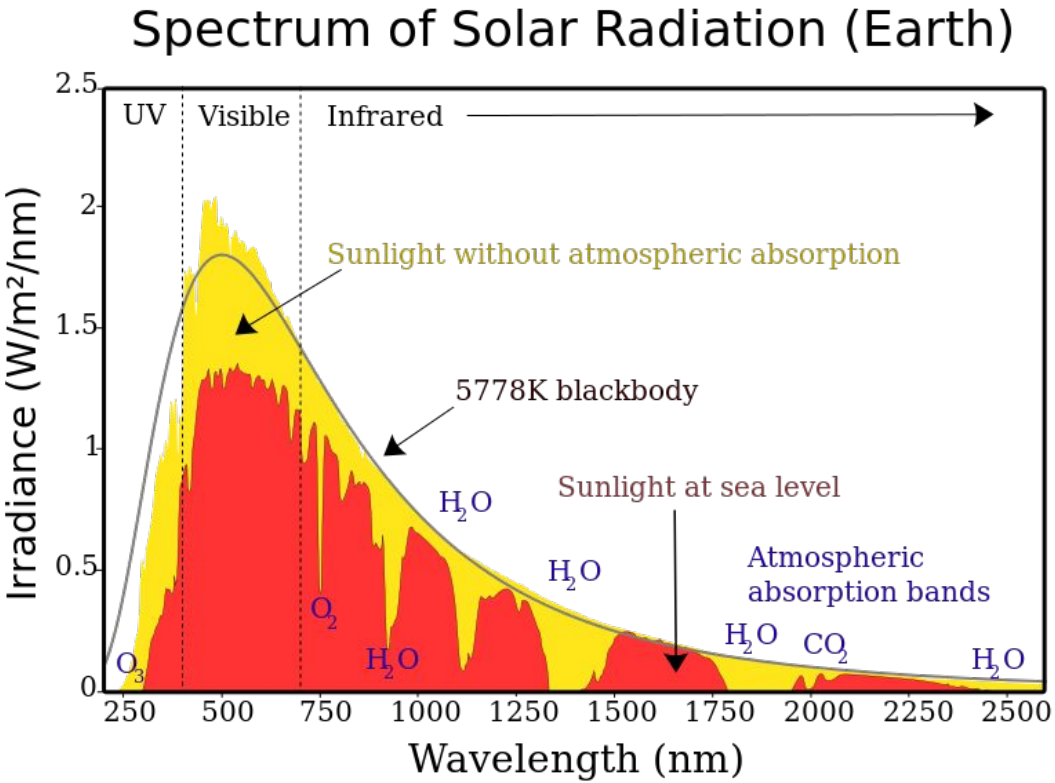
Assignment

Visual Encoding Using Color

Physics of Color



Physics of Color



What Not To Do With Color: Physics

Elevation is quantitative (-4000m to 1500m) ...

Wavelength is quantitative (400nm to 750nm)

So...

What Not To Do With Color: Physics

Elevation is quantitative (-4000m to 1500m) ...

Wavelength is quantitative (400nm to 750nm)

So...

elevation = -4000
λ = 400

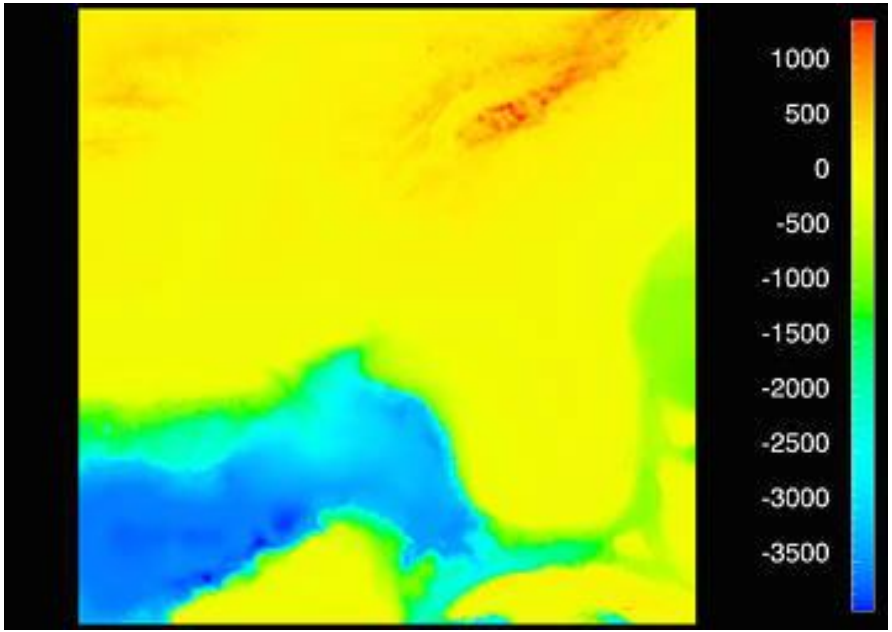
elevation = 1500
λ = 750

What Not To Do With Color: Physics

Elevation is quantitative (-4000m to 1500m) ...

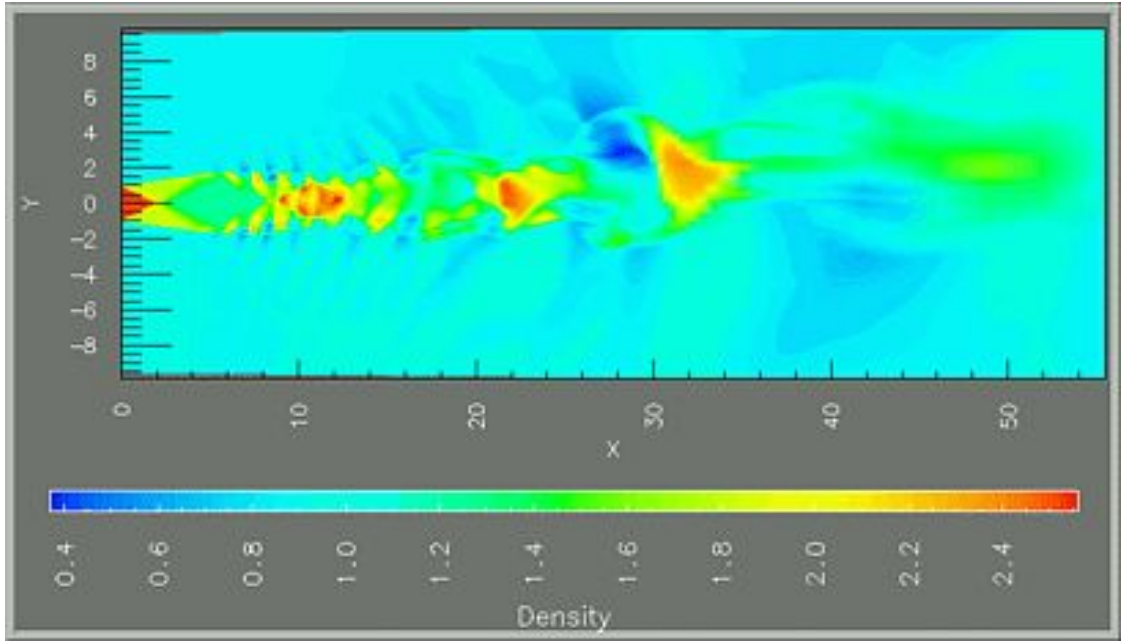
Wavelength is quantitative (400nm to 750nm)

So...



Jet / Rainbow Colormap Considered Harmful

Yet it’s the default behavior for matplotlib



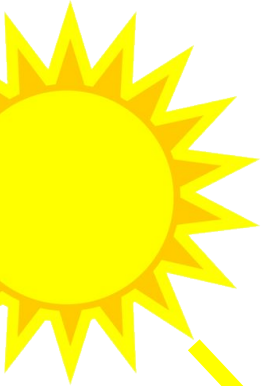
Perception

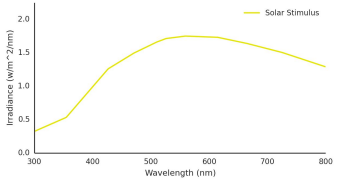
(it’s better than reality)

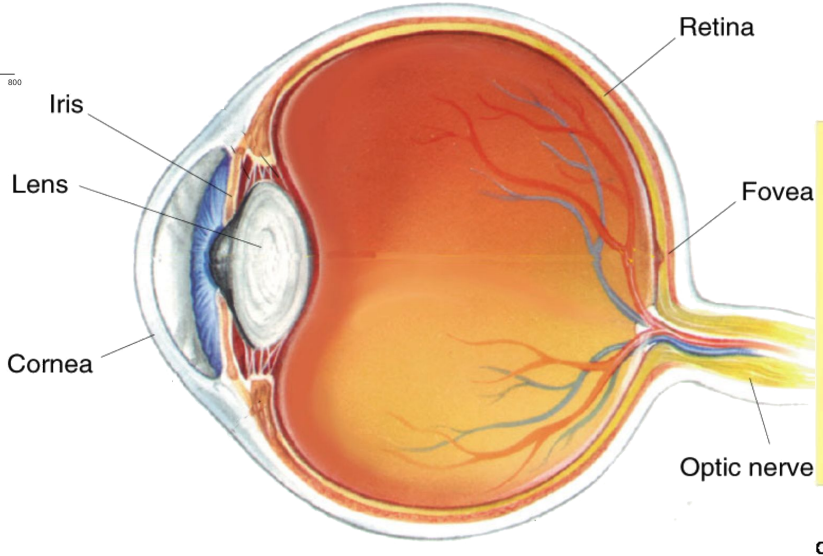
Human Color Vision:

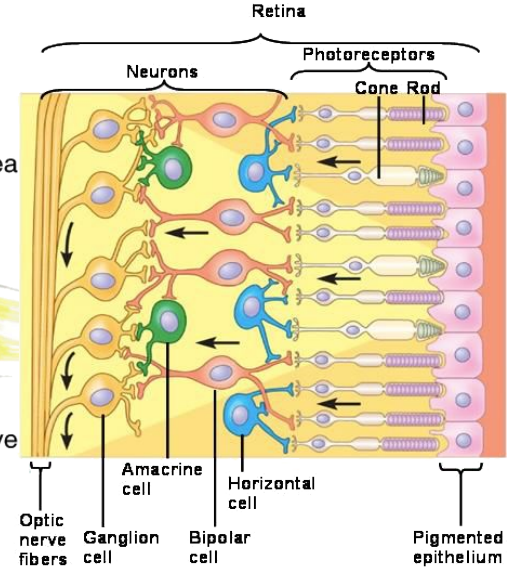
Tristimulus Theory

What To Do With Color: Perception

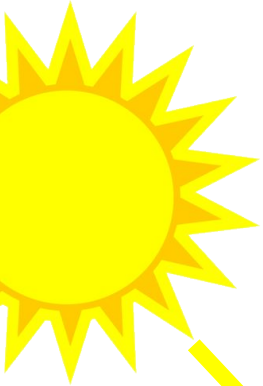


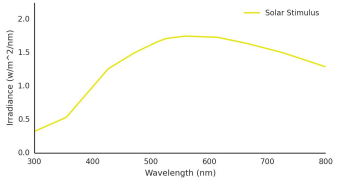


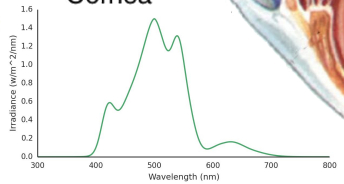


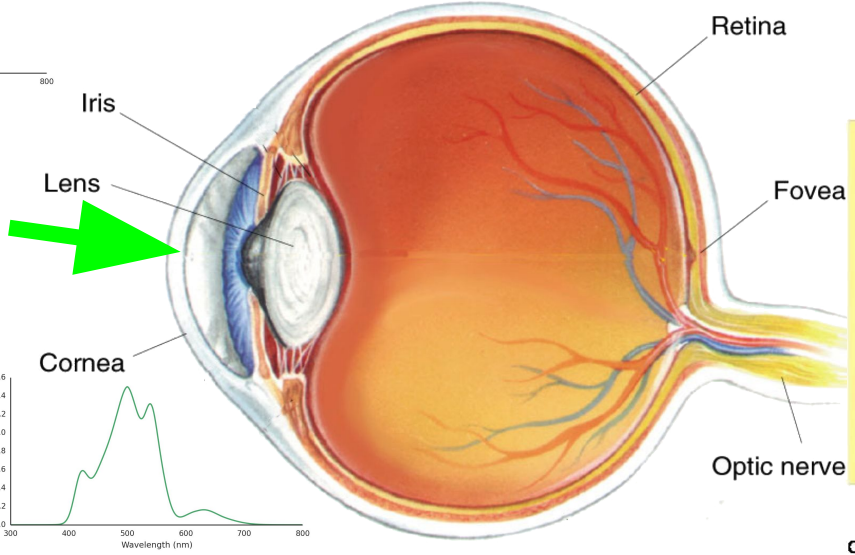


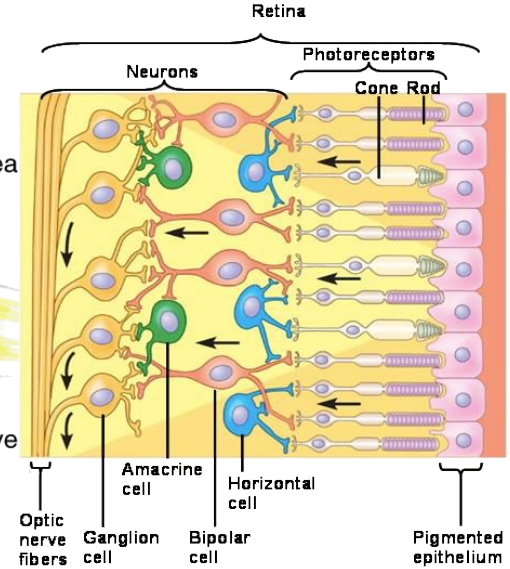
What To Do With Color: Perception



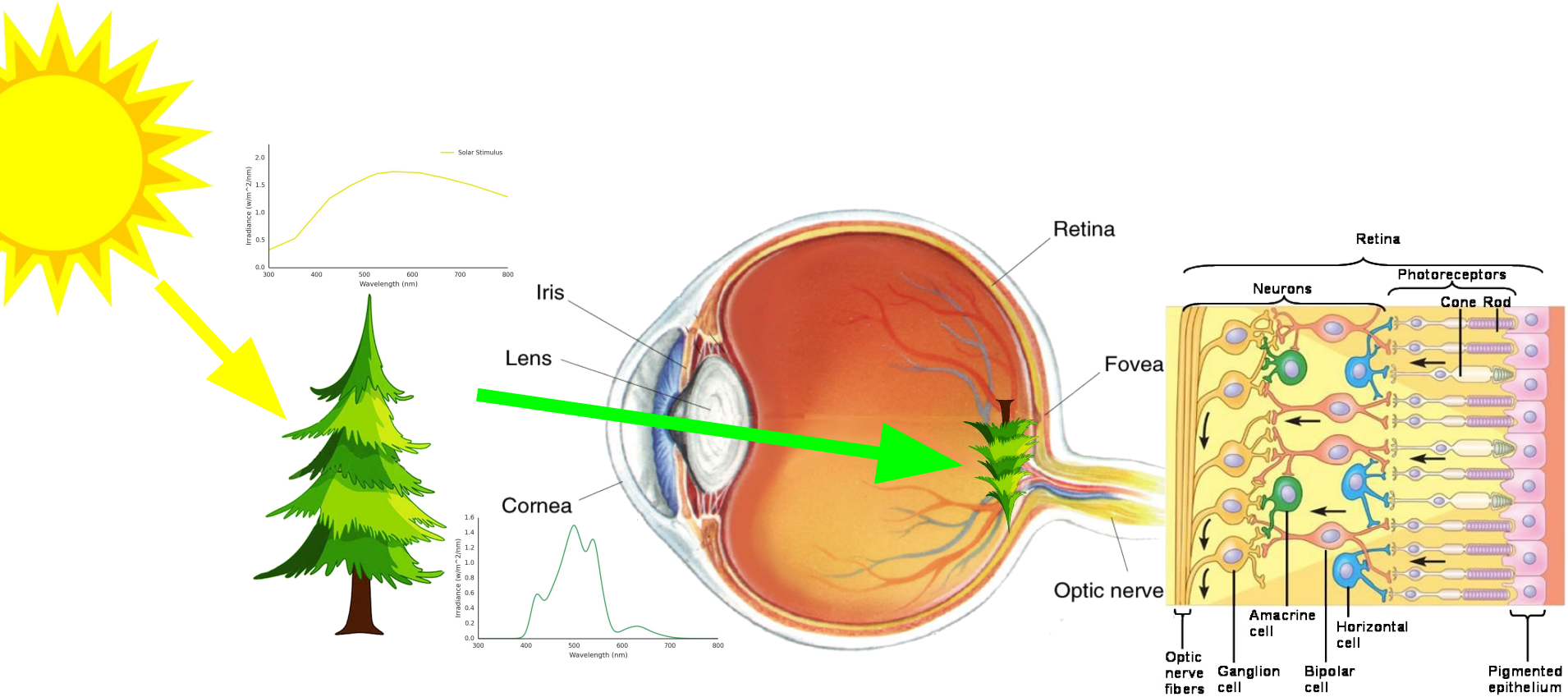


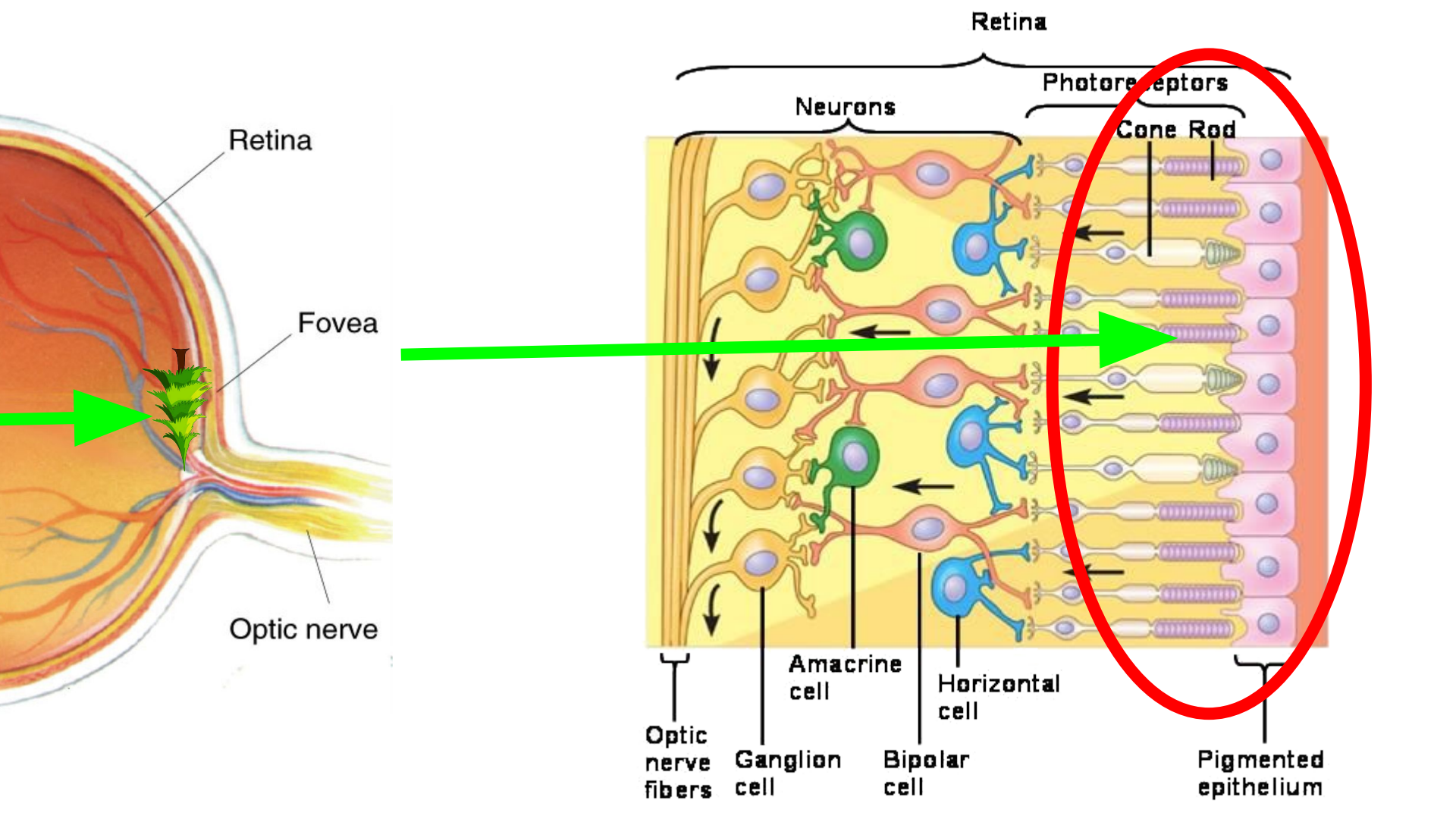


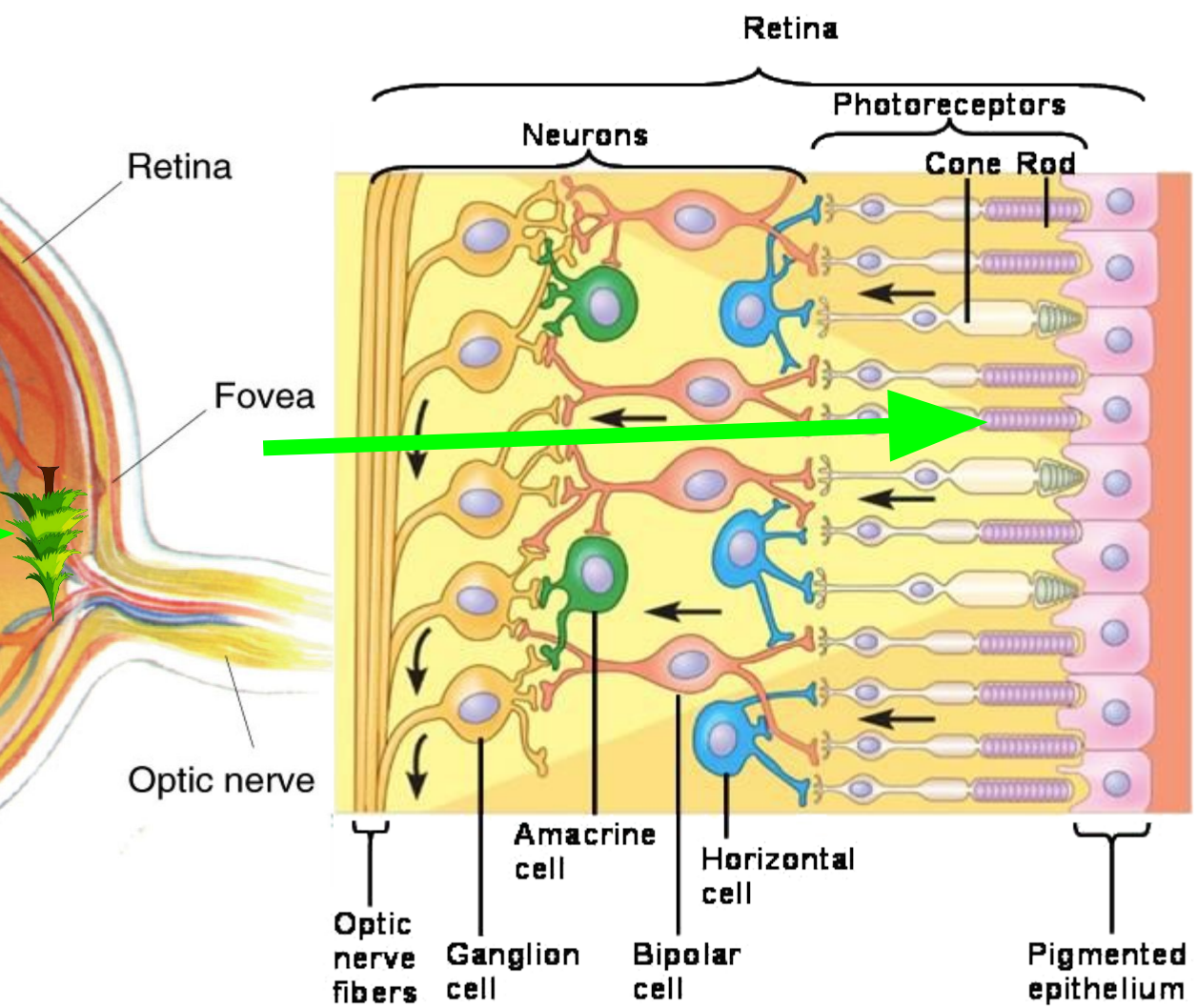


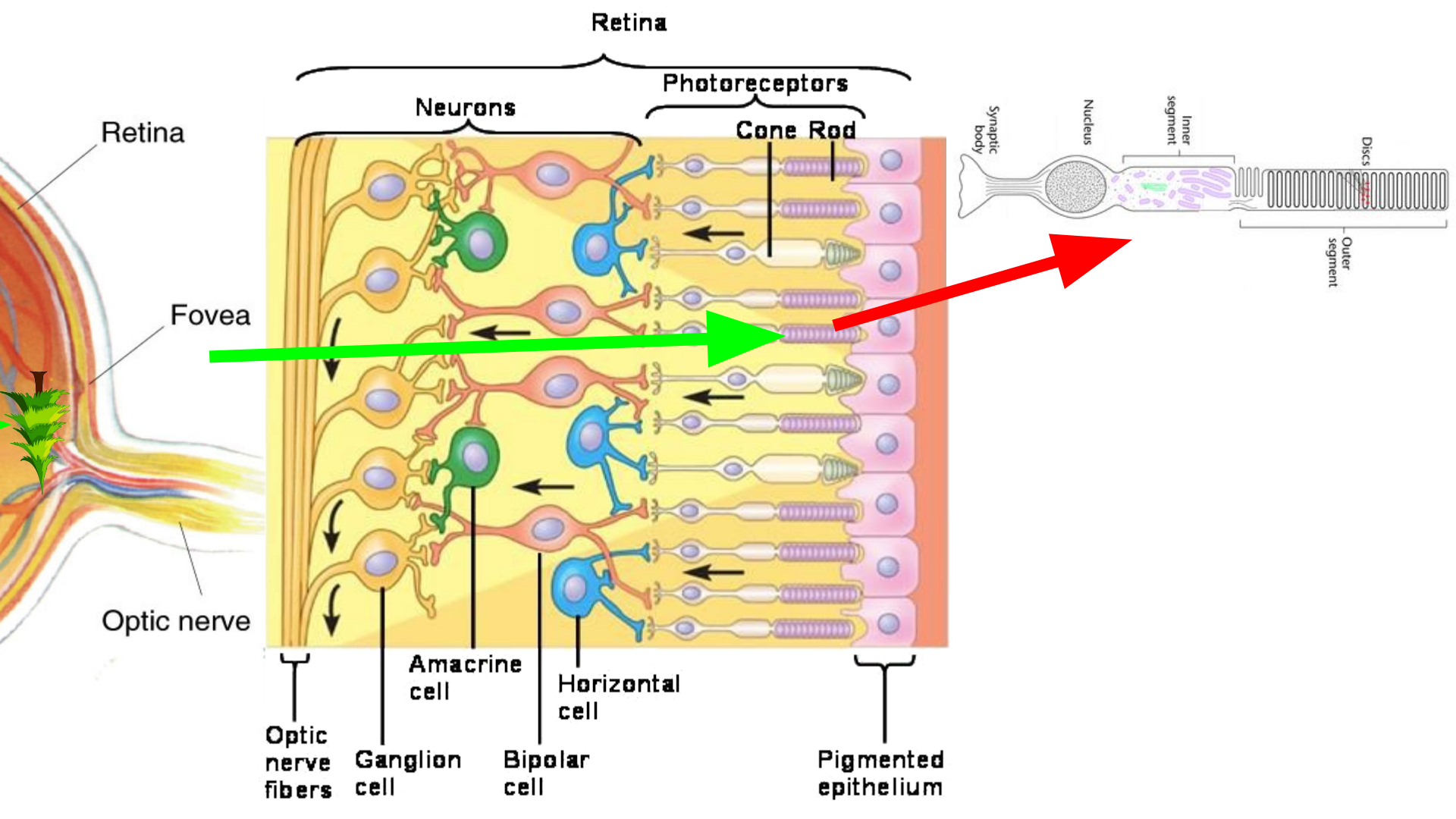


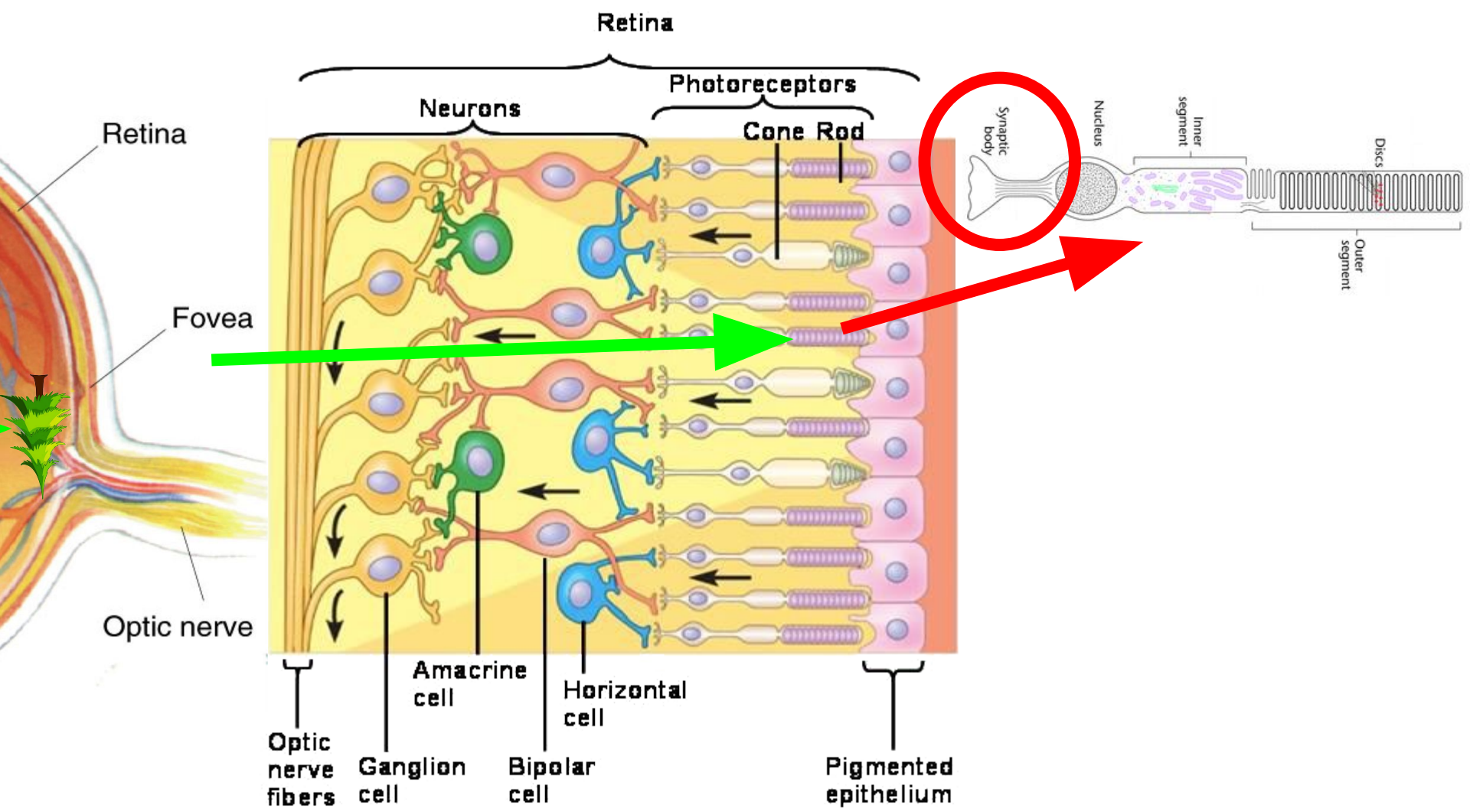
What To Do With Color: Perception

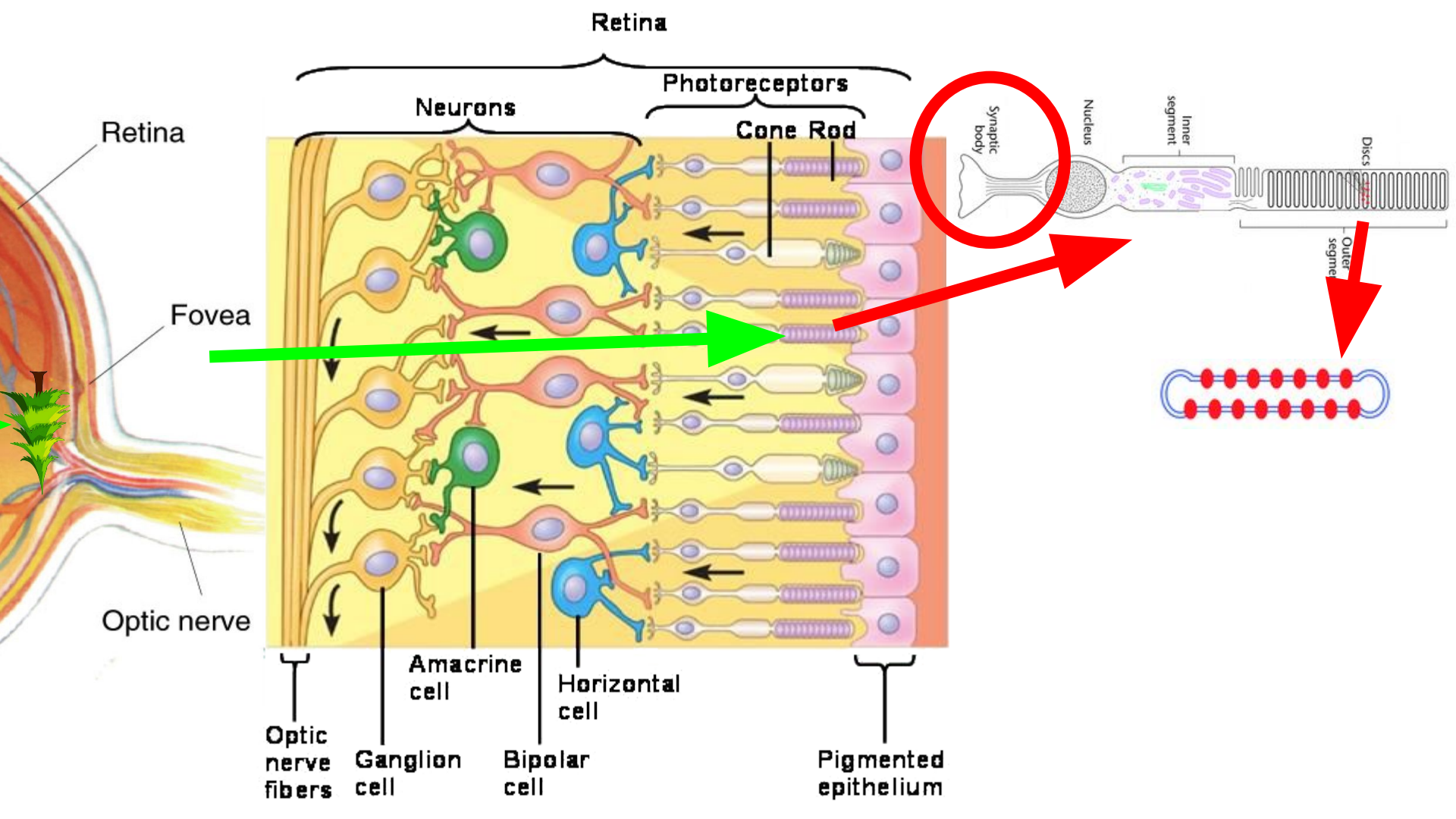


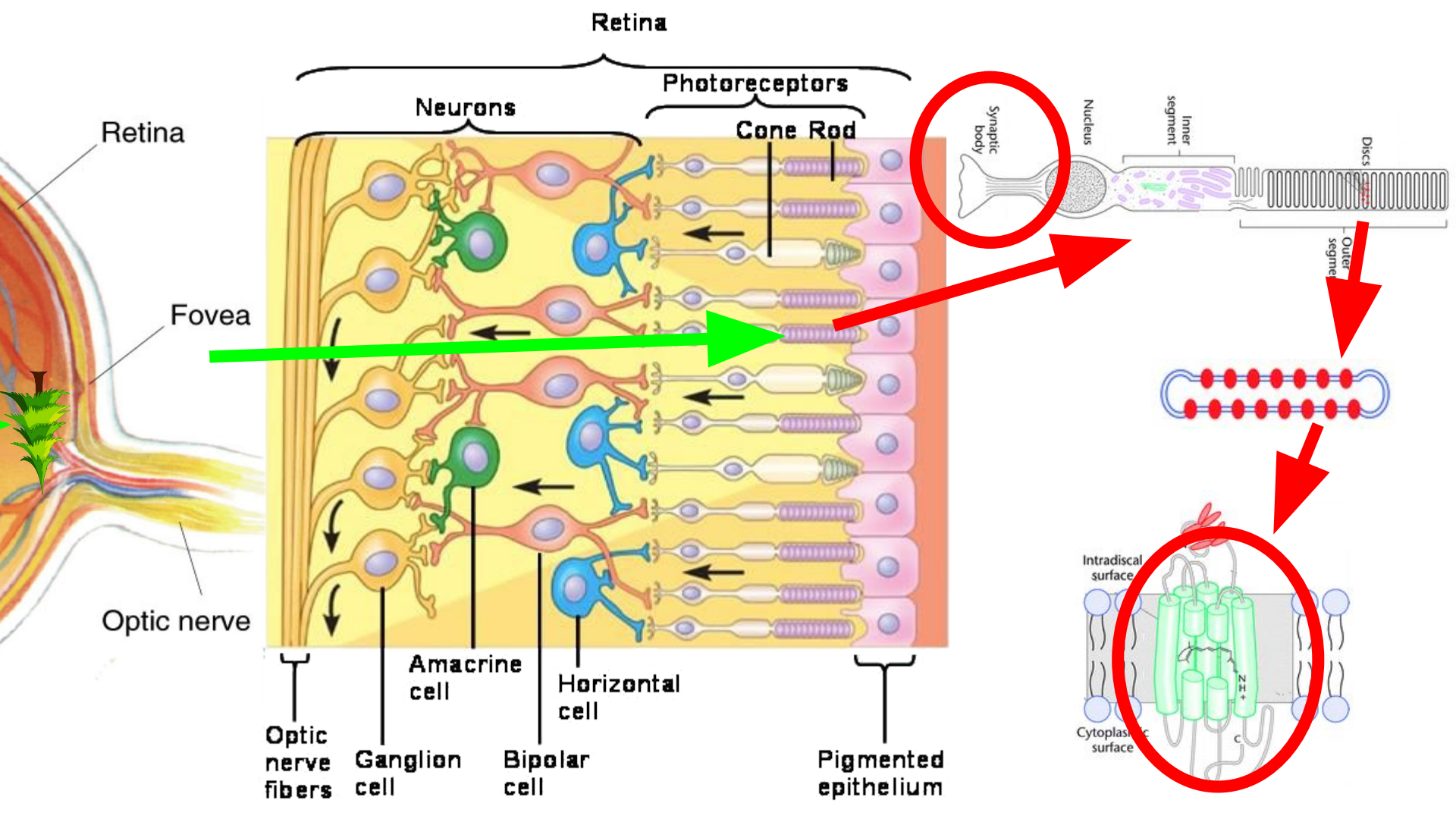

Retina
Fovea
Optic nerve
Retina
Photoreceptors
Neurons
Cone
Rod
Amacrine cell
Horizontal cell
Optic nerve fibers
Ganglion cell
Bipolar cell
Pigmented epithelium

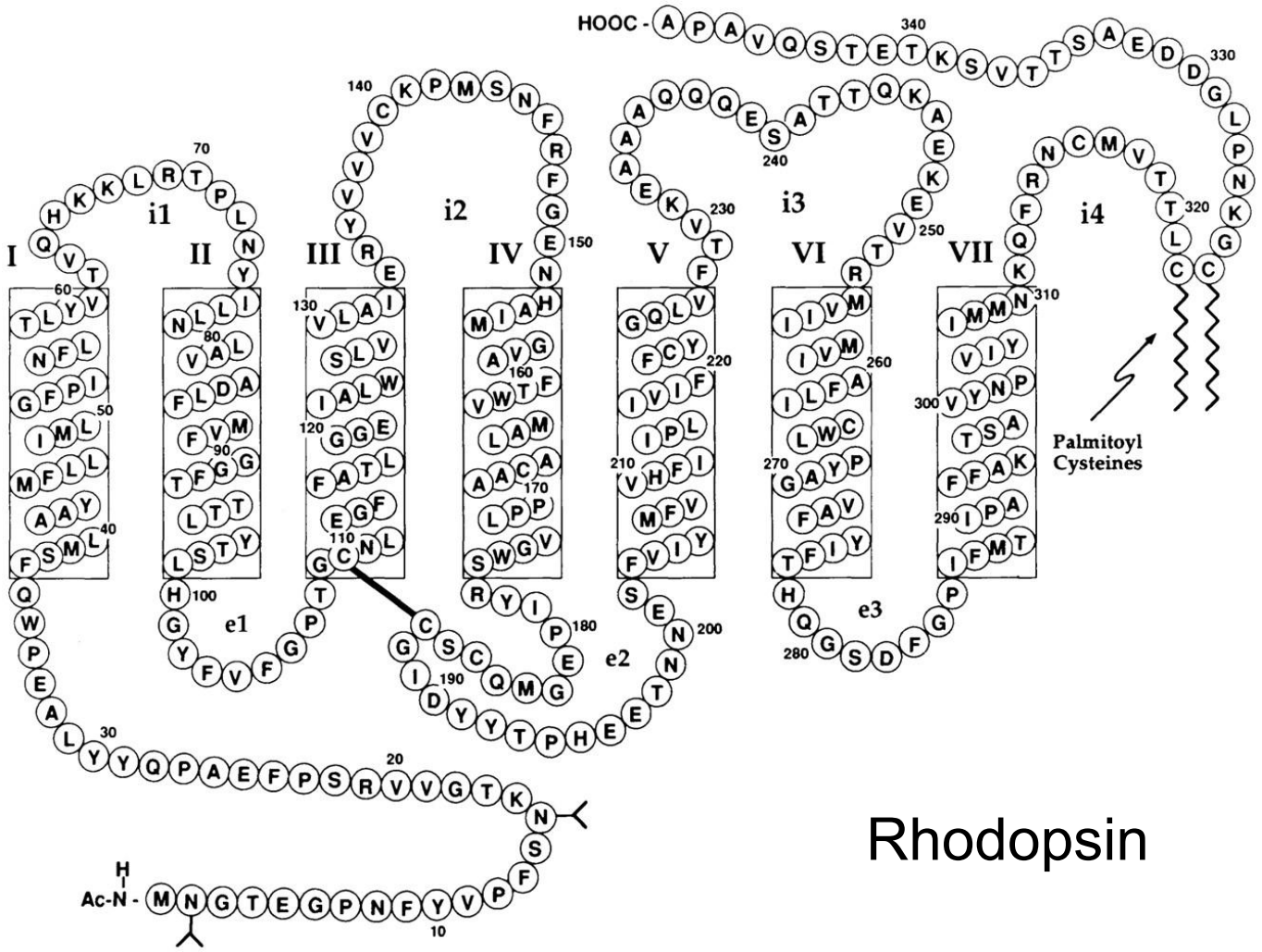

Retina
Retina
Neurons
Photoreceptors
Cone
Rod
Fovea
Optic nerve
Amacrine cell
Horizontal cell
Optic nerve fibers
Ganglion cell
Bipolar cell
Pigmented epithelium


Retina
Retina
Fovea
Optic nerve
Neurons
Photoreceptors
Cone
Rod
Optic nerve fibers
Ganglion cell
Amacrine cell
Bipolar cell
Horizontal cell
Pigmented epithelium
Synaptic body
Nucleus
Inner segment
Discs
Outer segment


Retina
Retina
Fovea
Optic nerve
Neurons
Photoreceptors
Cone
Rod
Amacrine cell
Horizontal cell
Optic nerve fibers
Ganglion cell
Bipolar cell
Pigmented epithelium
Synaptic body
Nucleus
Inner segment
Discs
Outer segment

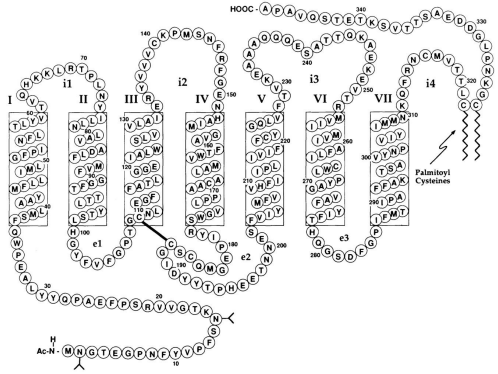

Retina
Retina
Fovea
Optic nerve
Neurons
Photoreceptors
Cone
Rod
Amacrine cell
Horizontal cell
Optic nerve fibers
Ganglion cell
Bipolar cell
Pigmented epithelium
Synaptic body
Nucleus
Inner segment
Discs
Outer segment

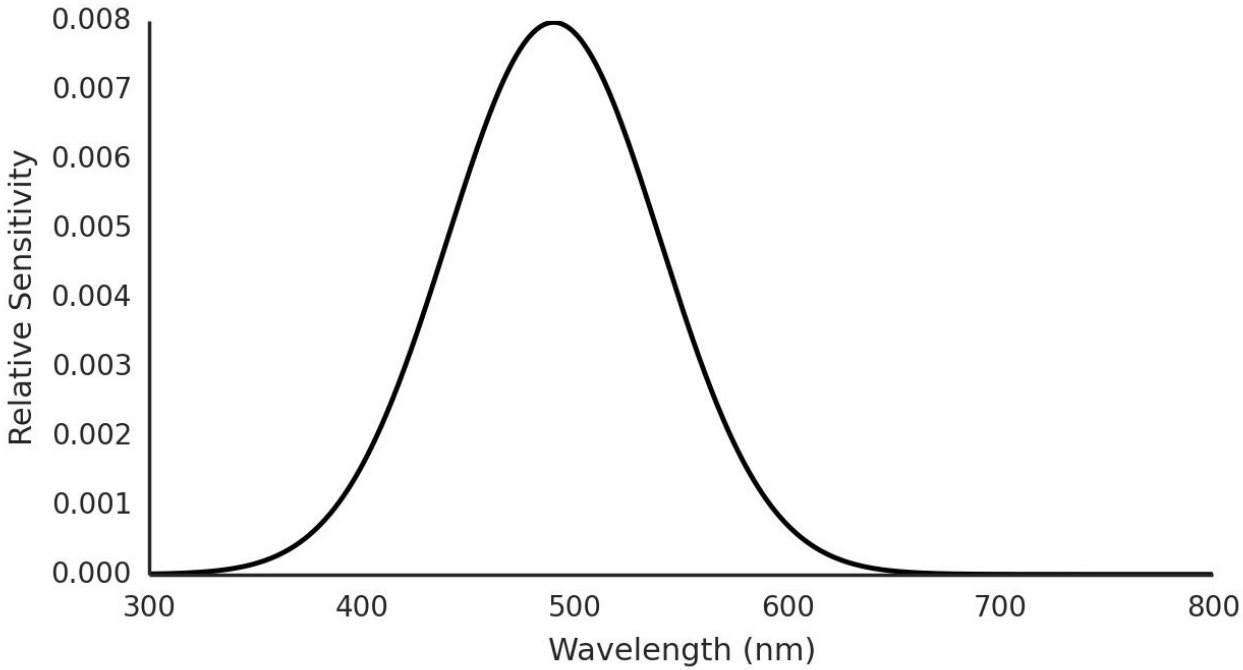

Retina
Retina
Fovea
Optic nerve
Neurons
Photoreceptors
Cone
Rod
Optic nerve fibers
Ganglion cell
Amacrine cell
Bipolar cell
Horizontal cell
Pigmented epithelium
Synaptic body
Nucleus
Inner segment
Discs
Outer segment
Intradiscal surface
Cytoplasmic surface

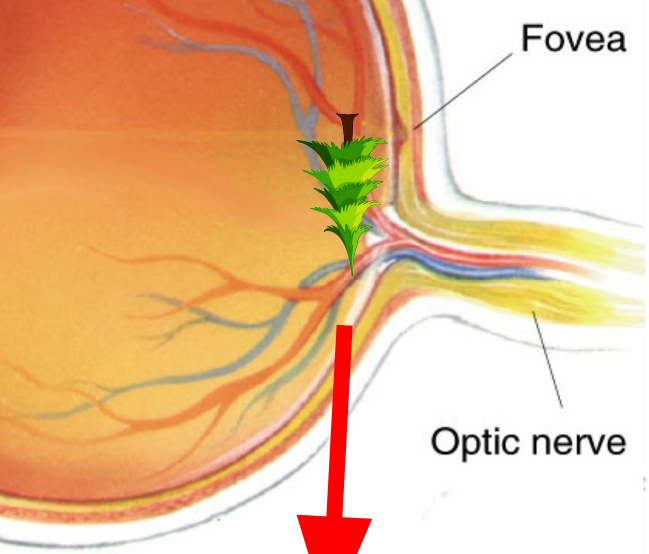

HOOC -
Ac-N -
H
I
II
III
IV
V
VI
VII
i1
i2
i3
i4
e1
e2
e3
Palmitoyl
Cysteines

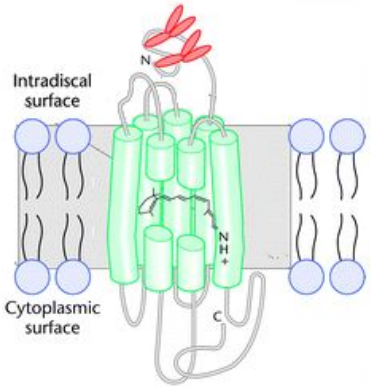
Rhodopsin

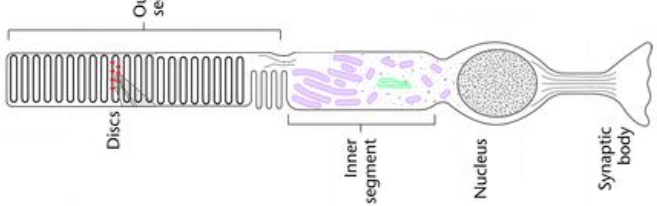
Rhodopsin Sensitivity

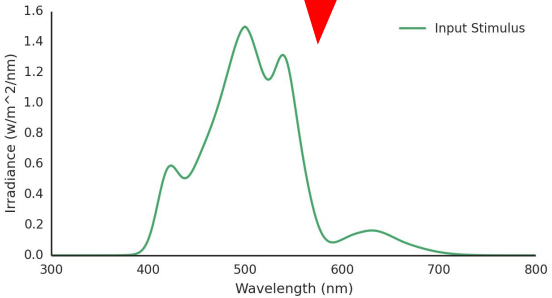




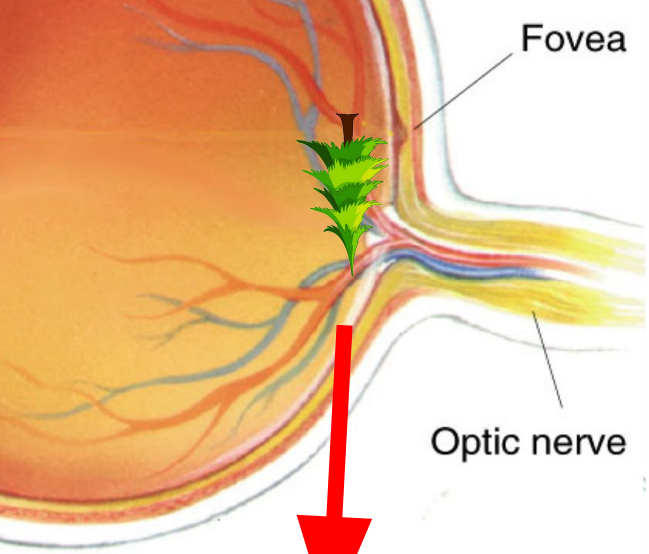

Fovea
Optic nerve

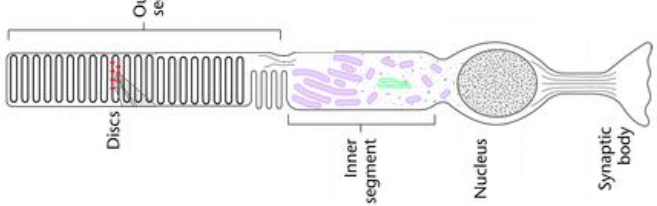

Intradiscal surface
Cytoplasmic surface
N
C
NH+

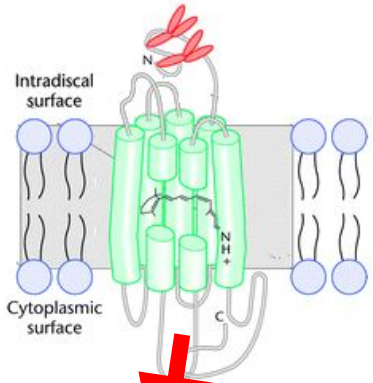

Discs
Inner segment
Nucleus
Synaptic body

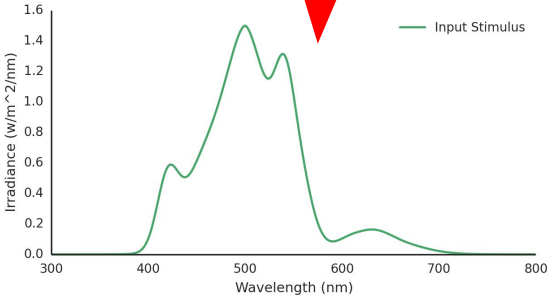

Input Stimulus
Irradiance (w/m^2/nm)
Wavelength (nm)
0.0
0.2
0.4
0.6
0.8
1.0
1.2
1.4
1.6
300
400
500
600
700
800

$$\Phi(\lambda)$$



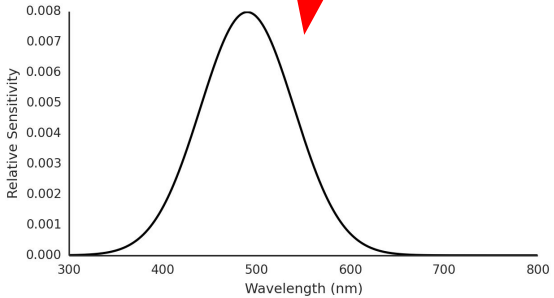






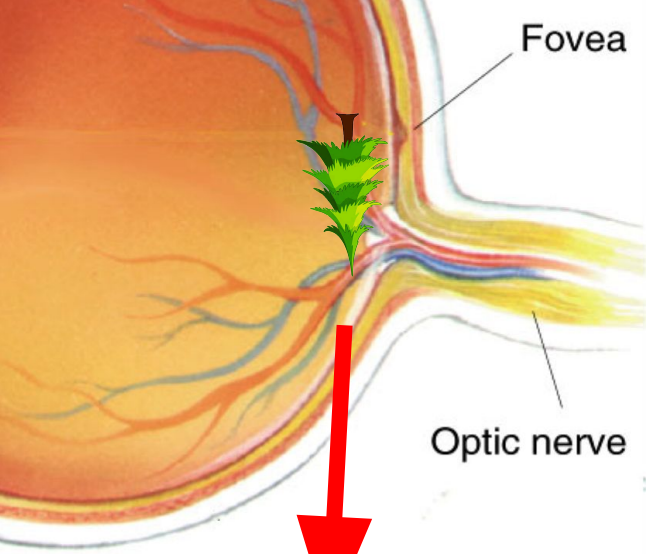
$\Phi(\lambda)$

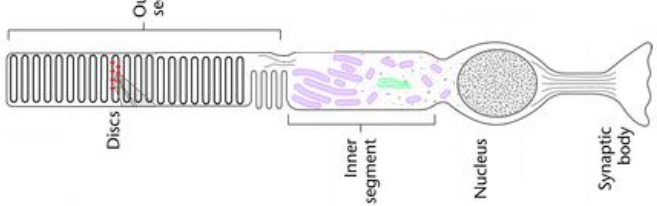
x

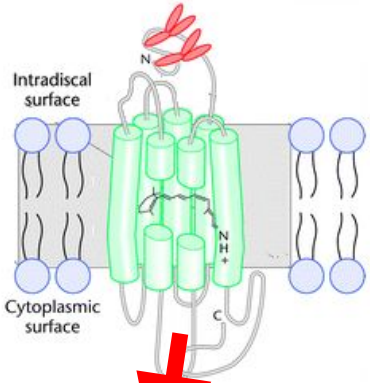


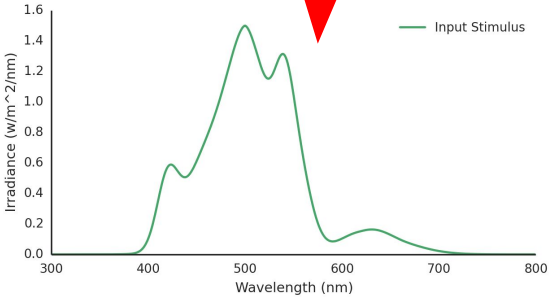
$R(\lambda)$

=



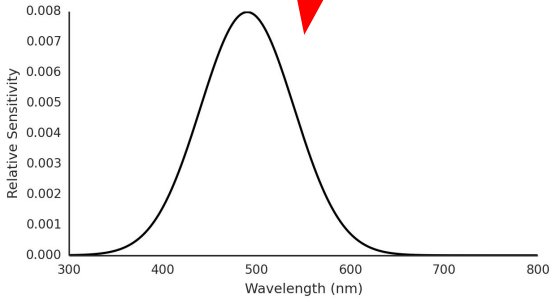






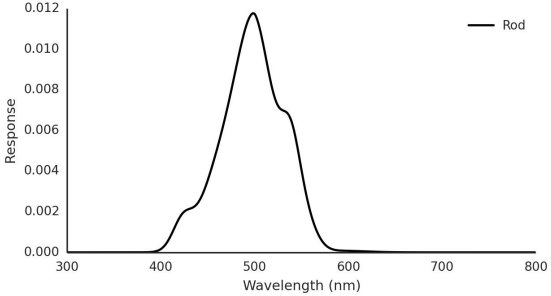
$\Phi(\lambda)$

X

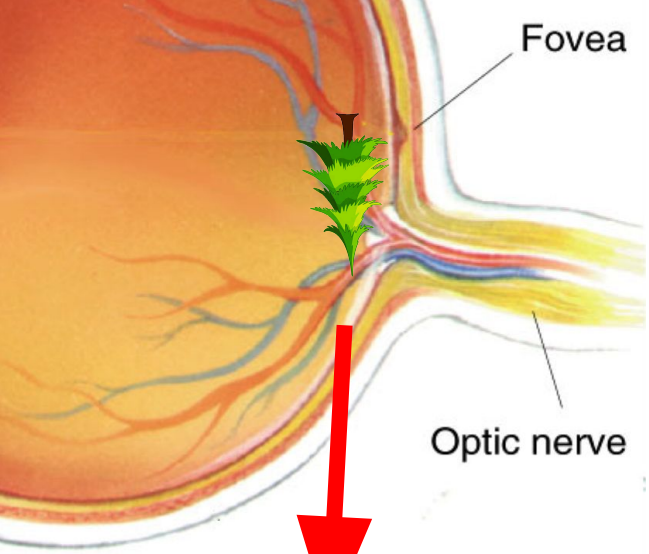


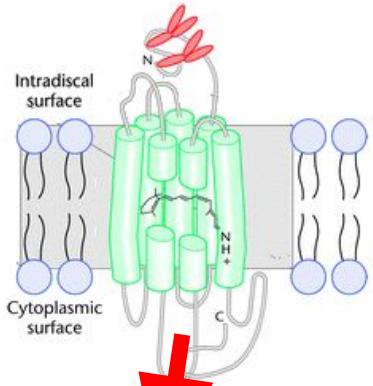
$R(\lambda)$

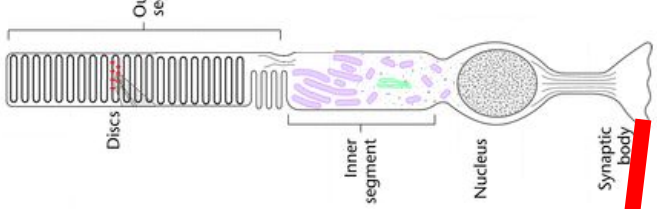
=

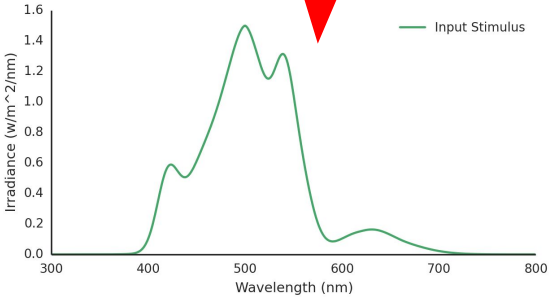


$R(\lambda)\Phi(\lambda)$



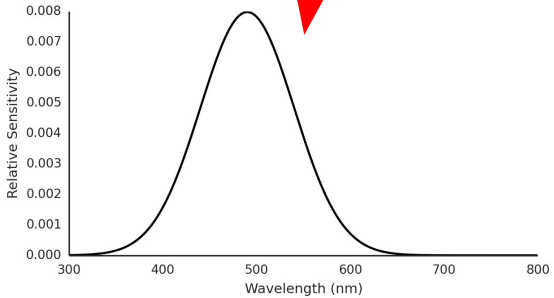






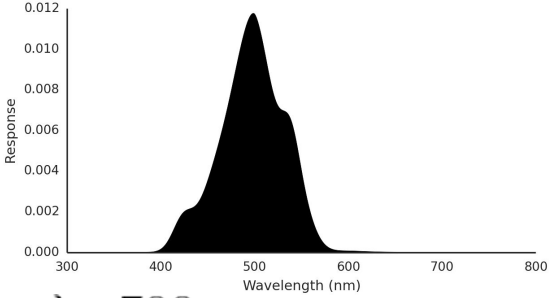
$\Phi(\lambda)$

x

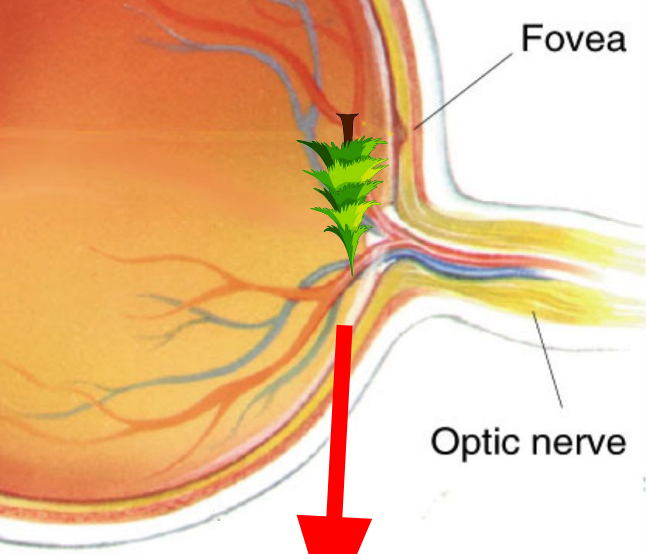


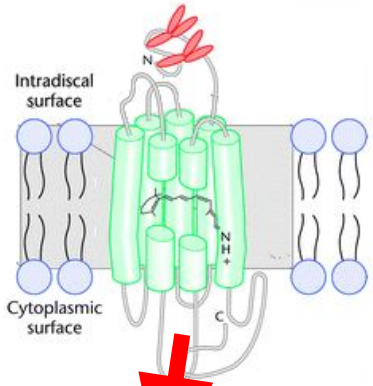
$R(\lambda)$

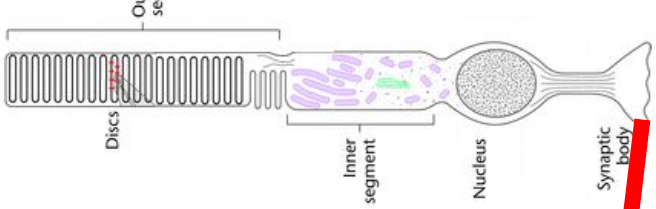
=

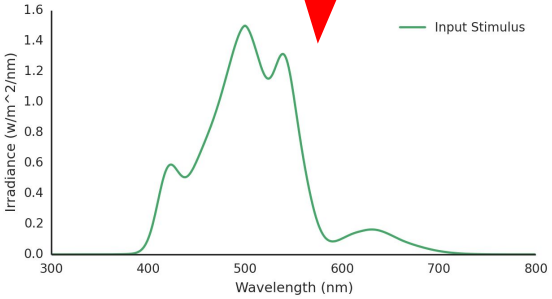


$$\int_{\lambda=380}^{\lambda=700} R(\lambda)\Phi(\lambda)\, d\lambda$$



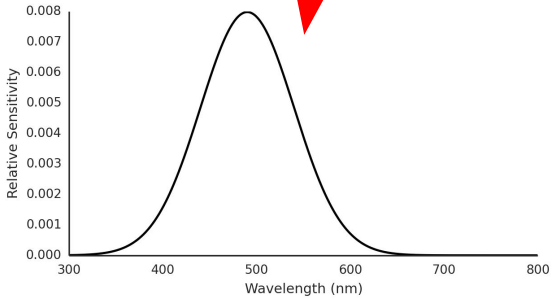






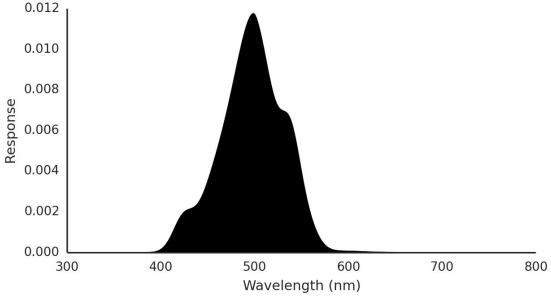
Φ

x

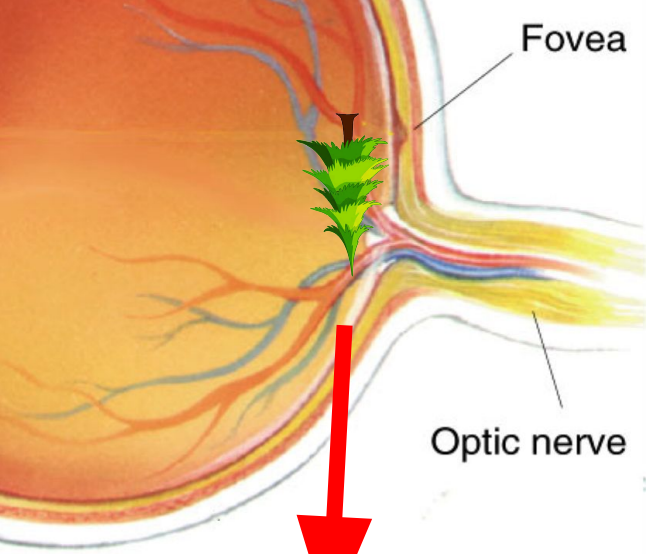


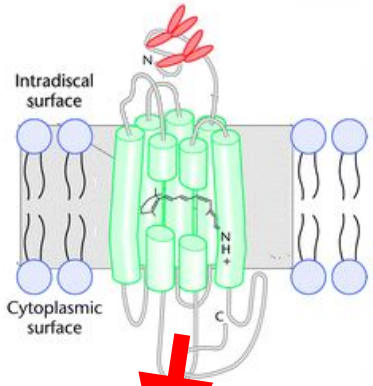
R

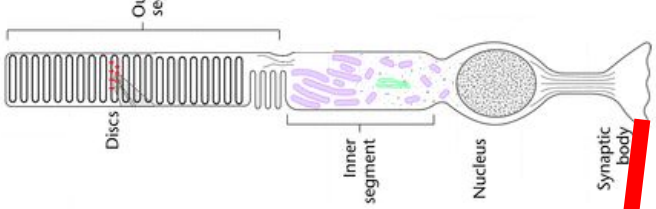
=

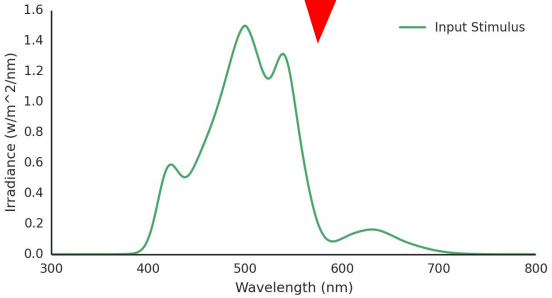


$$\int_{\Lambda} R\Phi$$



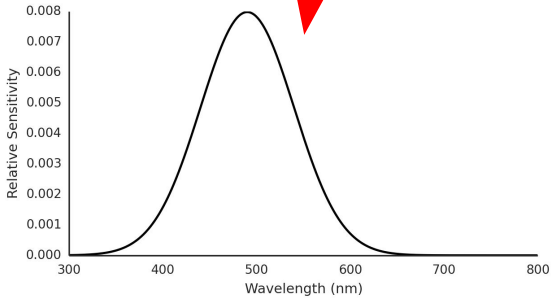






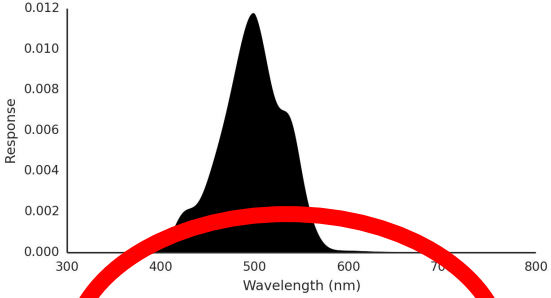
Φ

x



R

=



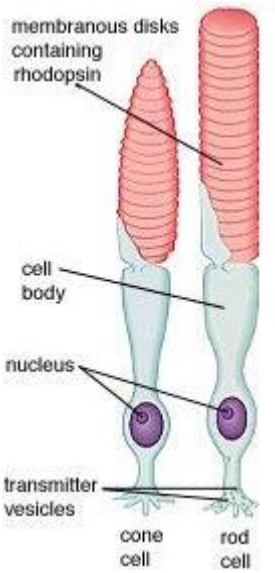
Stimulus Value

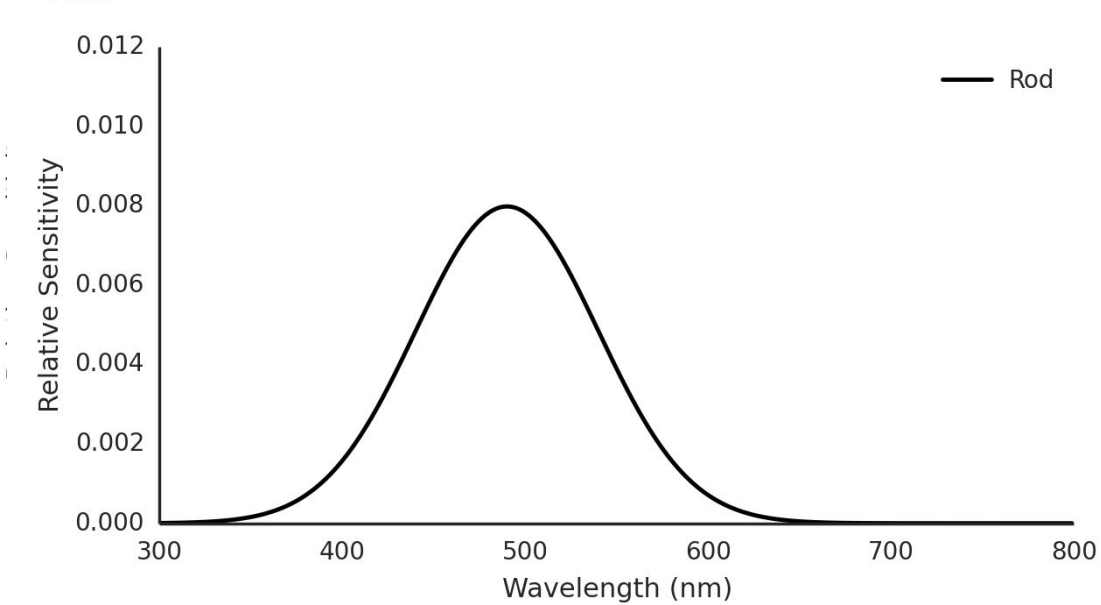
$$\int_{\Lambda} R\Phi$$

Photoreceptors

Rods (RHO): 490 nm

Chromosome 3



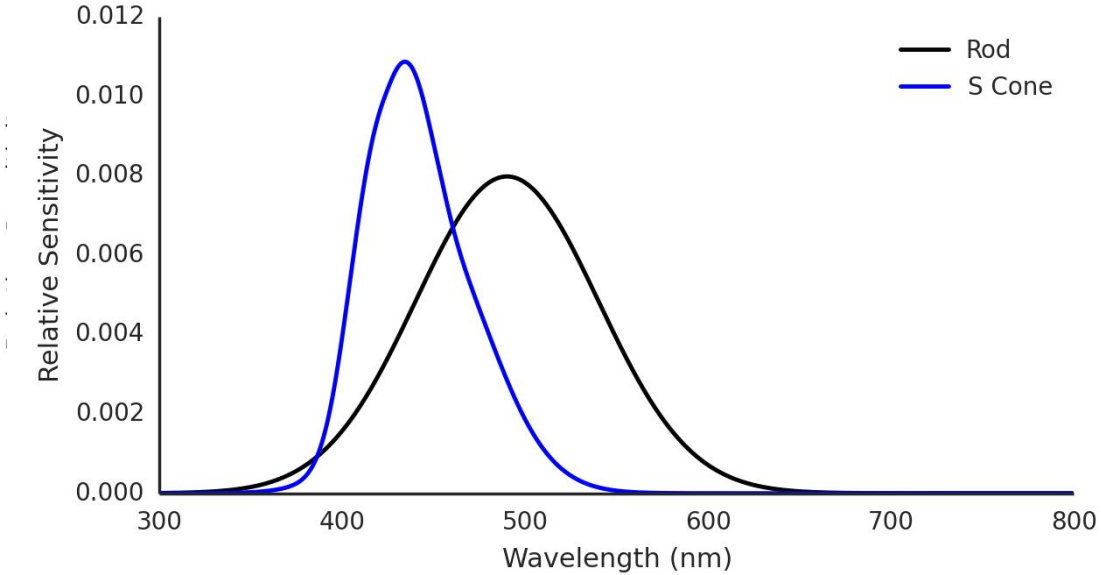


Photoreceptors

Rods (RHO): 490 nm
Chromosome 3

S cones (OPN1SW): 420 nm
Chromosome 7

S



Photoreceptors

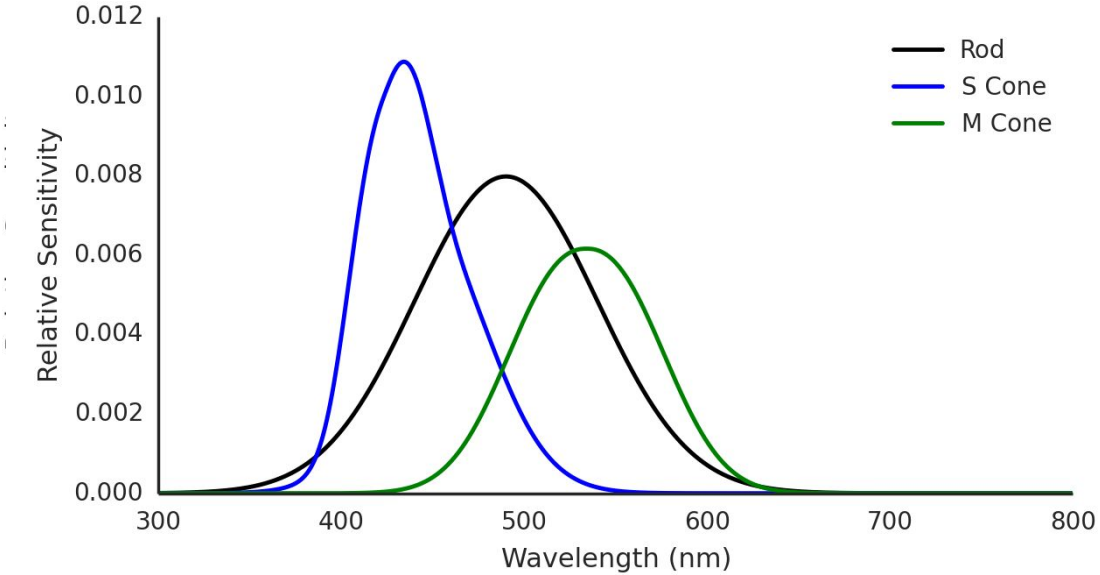
Rods (RHO): 490 nm
Chromosome 3

S cones (OPN1SW): 420 nm
Chromosome 7

M cones (OPN1MW): 534 nm
Chromosome 23 (X only)

S

M



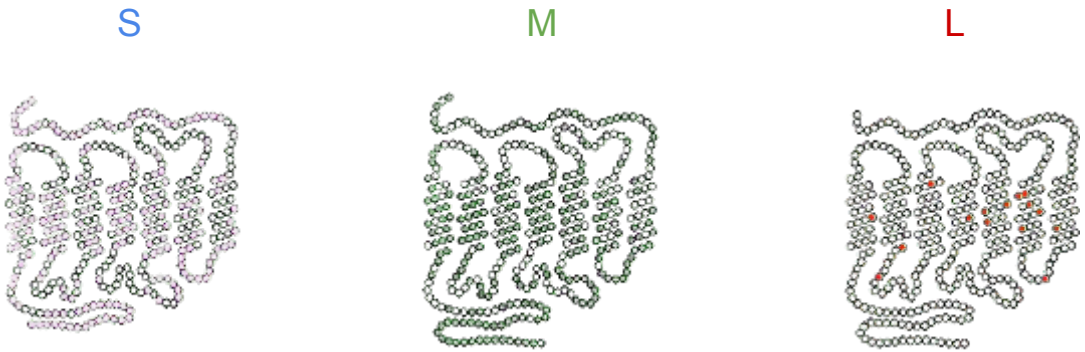
Photoreceptors

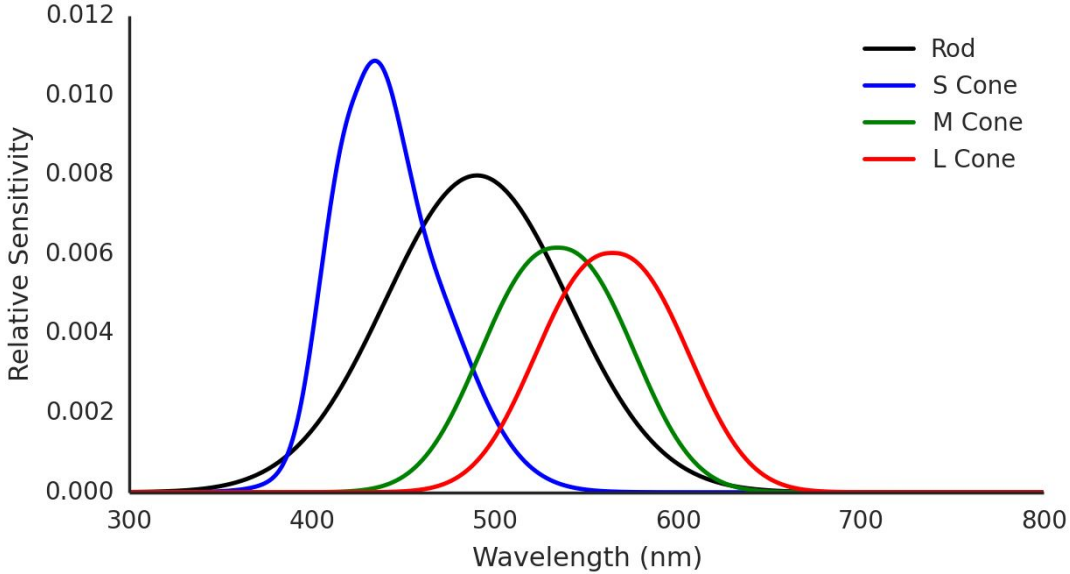
Rods (RHO): 490 nm
Chromosome 3

S cones (OPN1SW): 420 nm
Chromosome 7

M cones (OPN1MW): 534 nm
Chromosome 23 (X only)

L cones (OPN1LW): 564 nm
Chromosome 23 (X only)





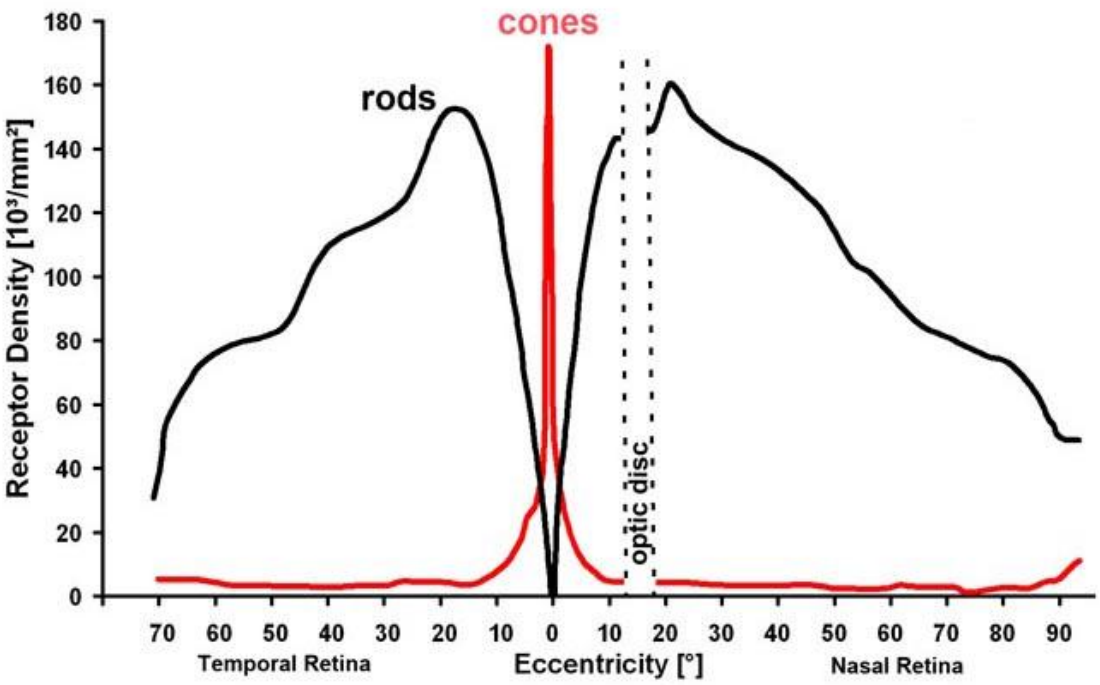
Rod and Cone Domains

Cones

- High light intensity
- Near the fovea

Rods:

- Low light intensity
- Away from the fovea

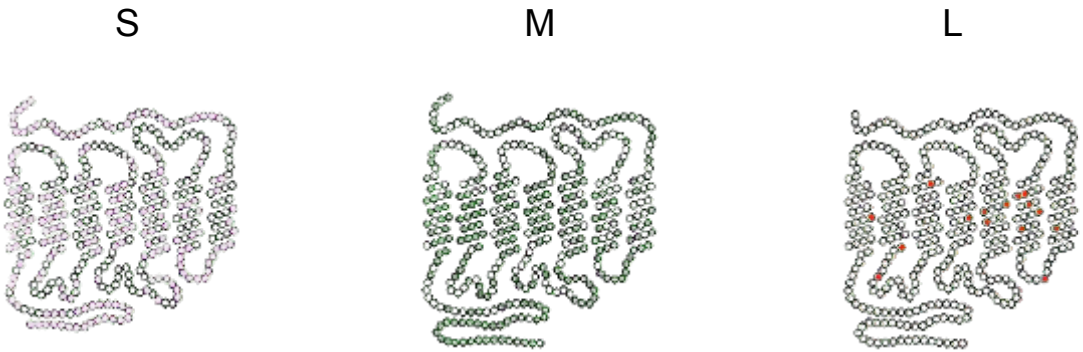


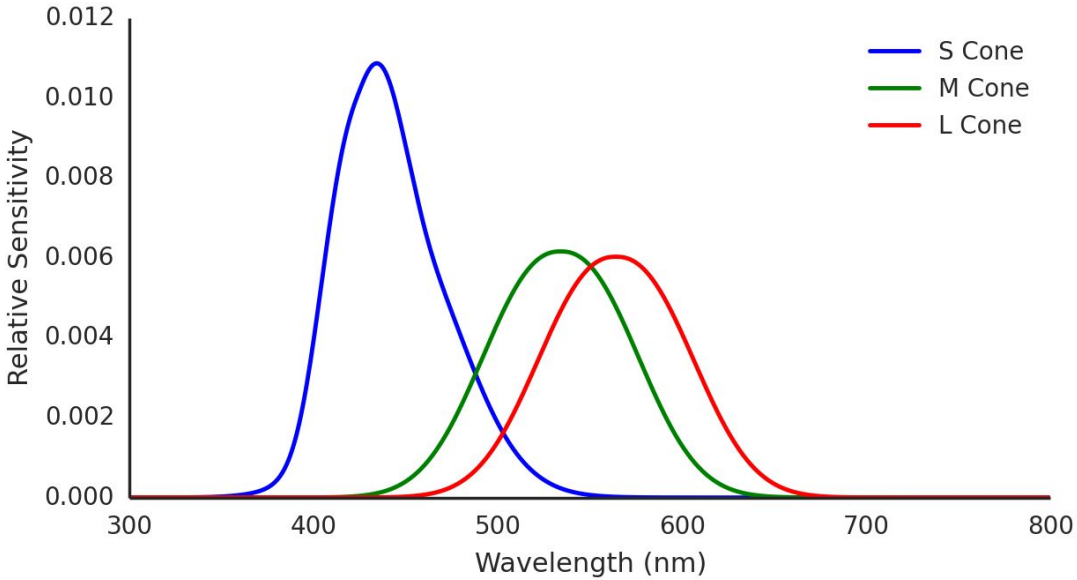
Cone Stimulation

S cones: 420 nm
Chromosome 7

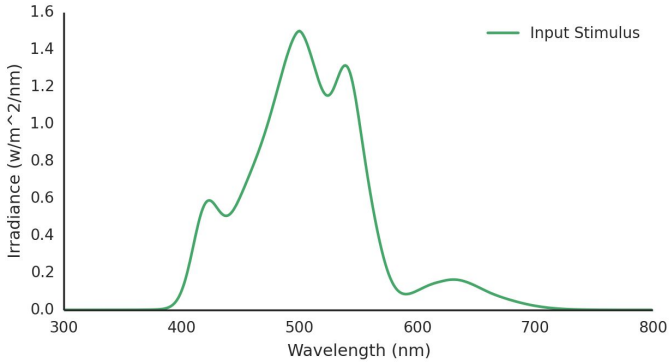
M cones: 534 nm
Chromosome 23 (X only)

L cones: 564 nm
Chromosome 23 (X only)

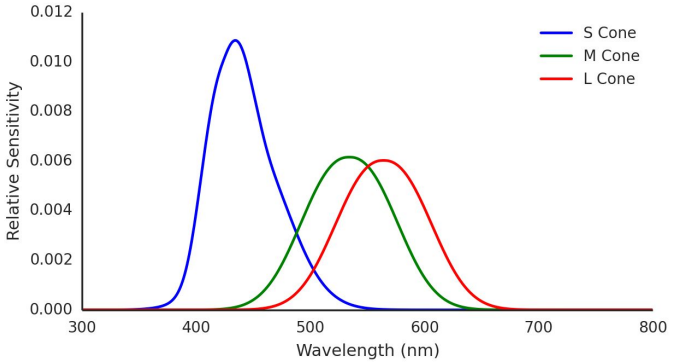




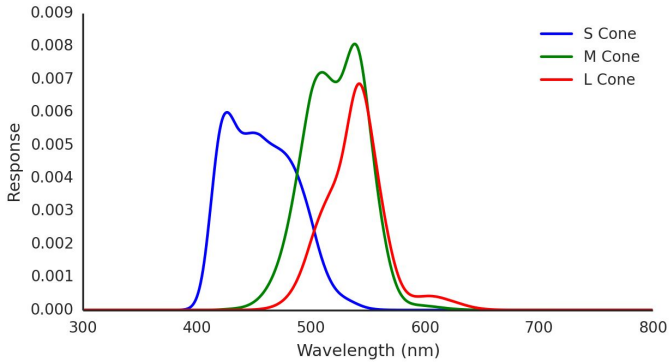
Tristimulus (LMS) Values



x



=

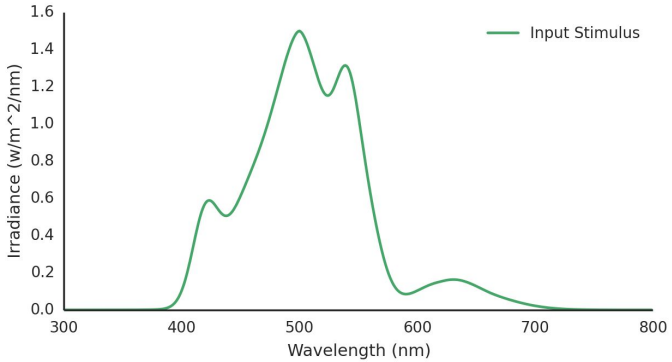


$$S \text{ stimulus} = \int_{\Lambda} S\Phi$$

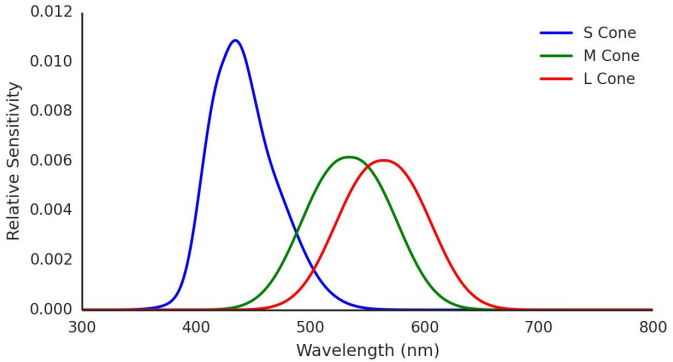
$$M \text{ stimulus} = \int_{\Lambda} M\Phi$$

$$L \text{ stimulus} = \int_{\Lambda} L\Phi$$

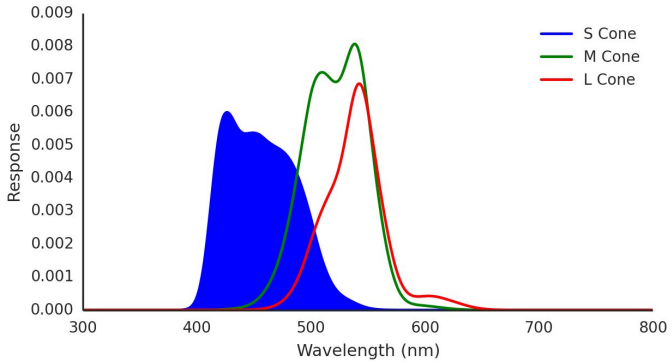
Tristimulus (LMS) Values



x



=

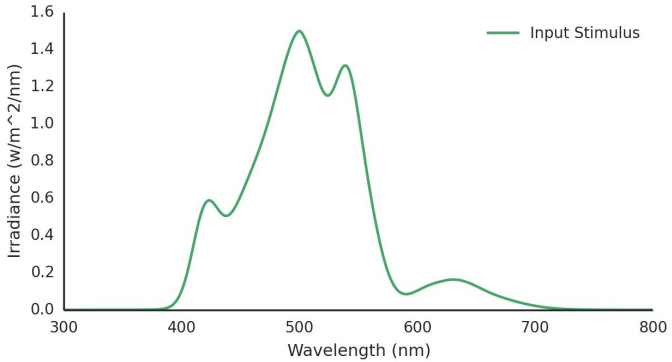


$$S \text{ stimulus} = \int_\Lambda S\Phi$$

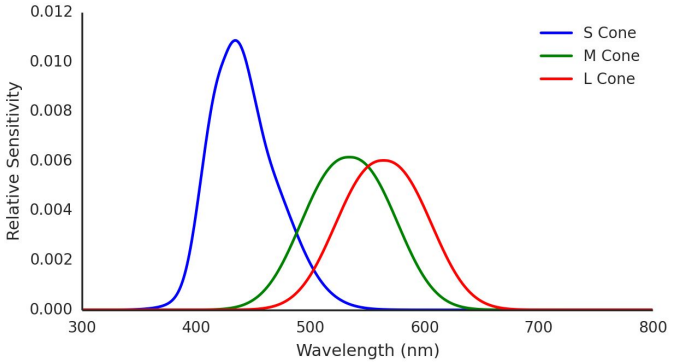
$$M \text{ stimulus} = \int_\Lambda M\Phi$$

$$L \text{ stimulus} = \int_\Lambda L\Phi$$

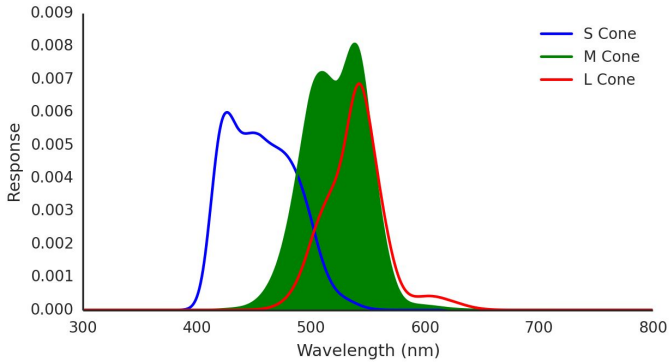
Tristimulus (LMS) Values



x



=

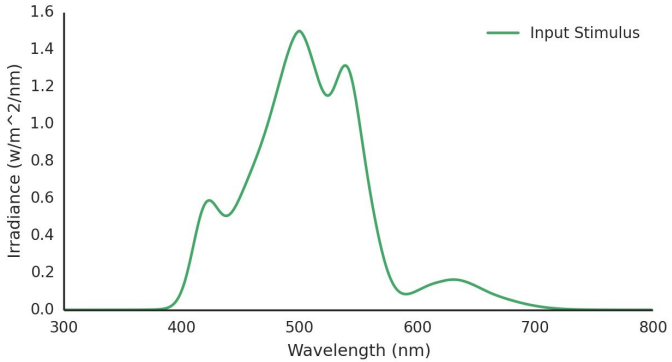


$$S \text{ stimulus} = \int_\Lambda S\Phi$$

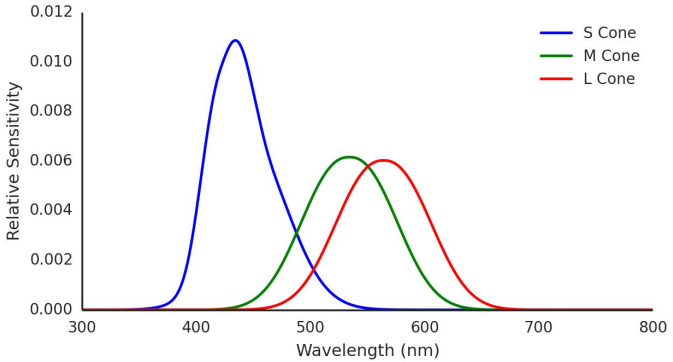
$$M \text{ stimulus} = \int_\Lambda M\Phi$$

$$L \text{ stimulus} = \int_\Lambda L\Phi$$

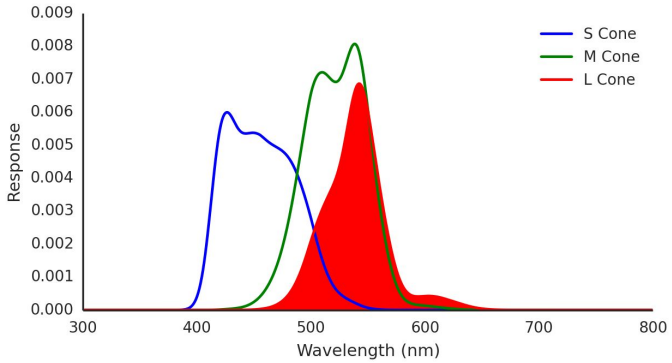
Tristimulus (LMS) Values



x



=

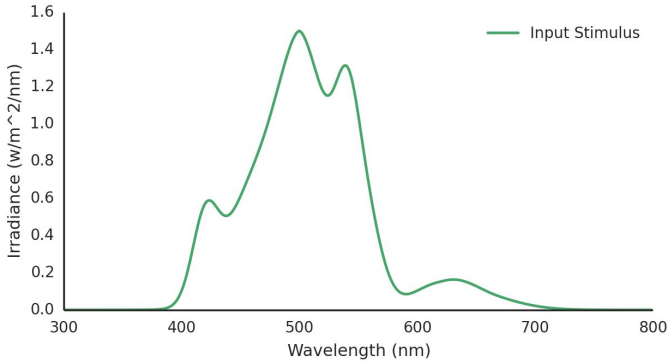


$$S \text{ stimulus} = \int_\Lambda S\Phi$$

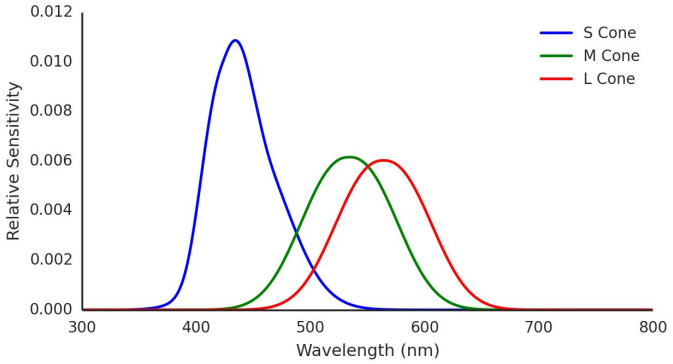
$$M \text{ stimulus} = \int_\Lambda M\Phi$$

$$L \text{ stimulus} = \int_\Lambda L\Phi$$

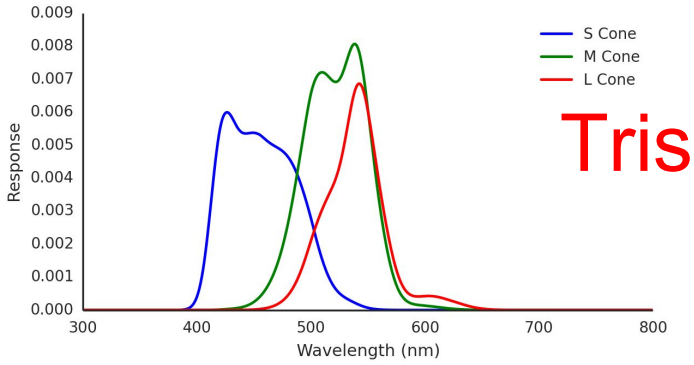
Tristimulus (LMS) Values



x



=



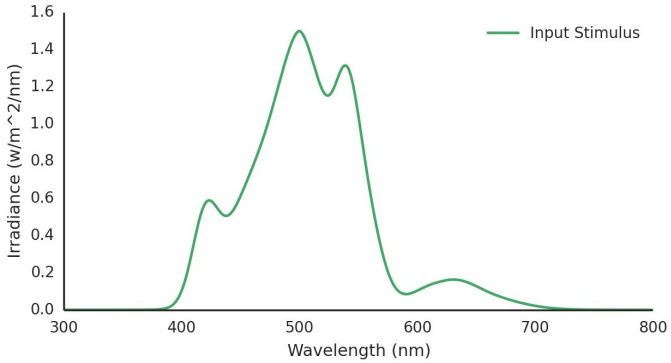
Tristimulus Value

$$S \text{ stimulus} = \int_\Lambda S\Phi$$

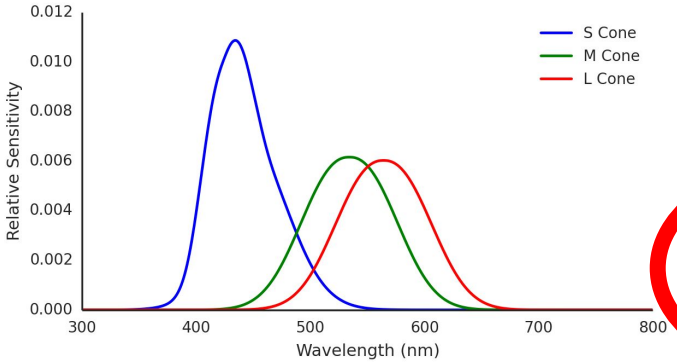
$$M \text{ stimulus} = \int_\Lambda M\Phi$$

$$L \text{ stimulus} = \int_\Lambda L\Phi$$

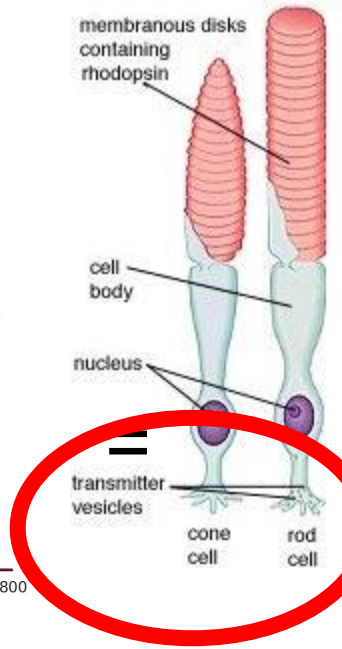
Tristimulus (LMS) Values

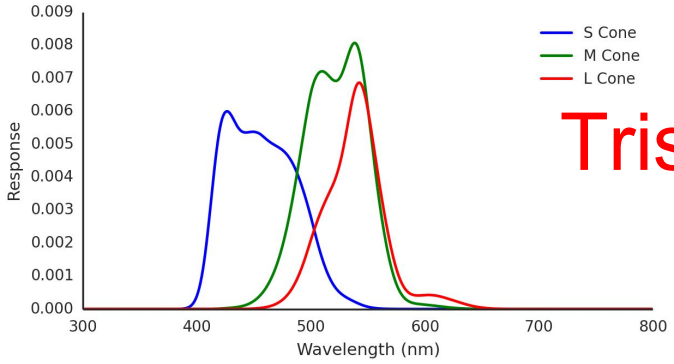


x



=



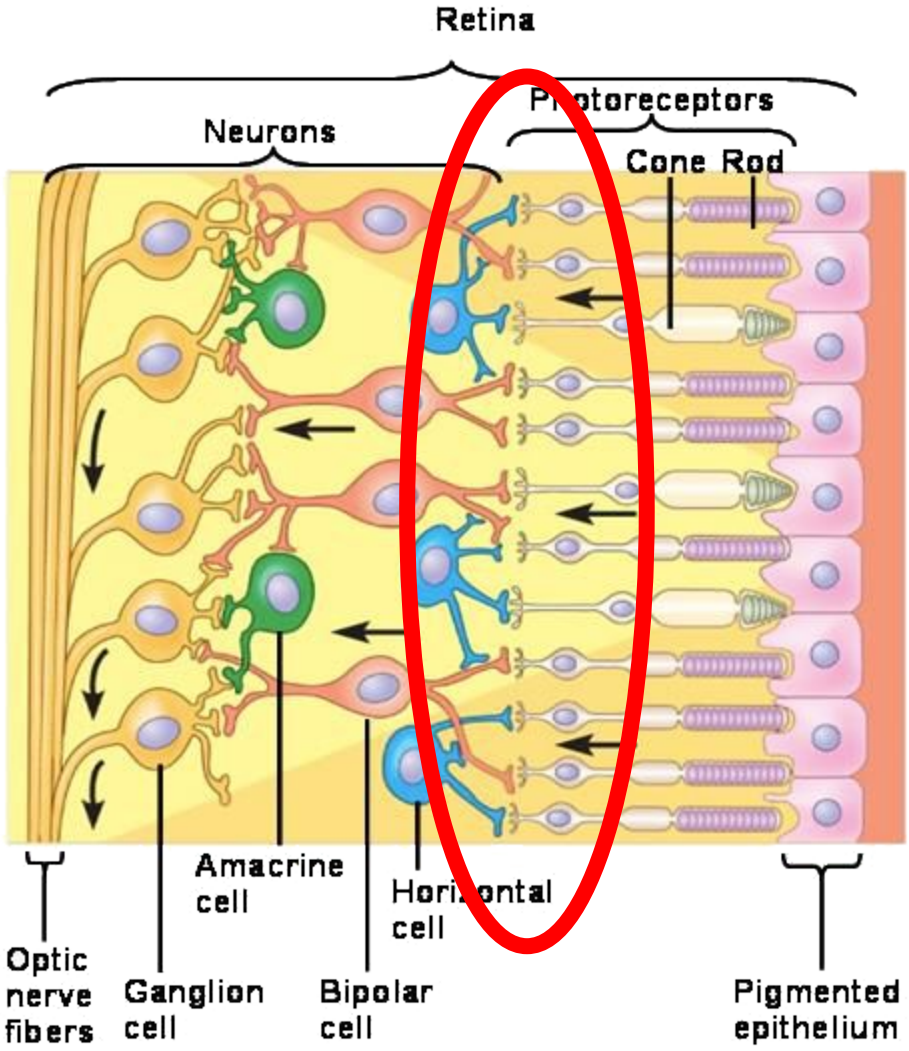


Tristimulus Value

$$S \text{ stimulus} = \int_\Lambda S\Phi$$

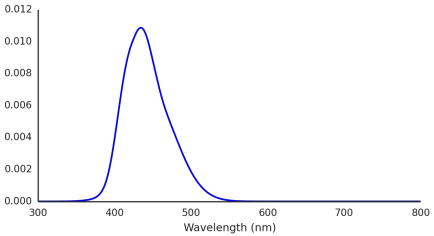
$$M \text{ stimulus} = \int_\Lambda M\Phi$$

$$L \text{ stimulus} = \int_\Lambda L\Phi$$


Retina
Neurons
Photoreceptors
Cone
Rod
Amacrine cell
Horizontal cell
Optic nerve fibers
Ganglion cell
Bipolar cell
Pigmented epithelium

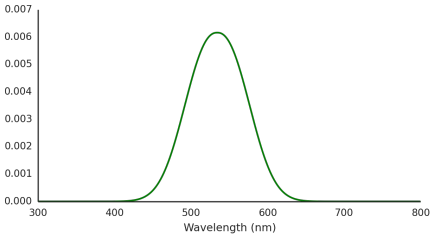
Colorblindness

EVERTON LOPES
©2014



X

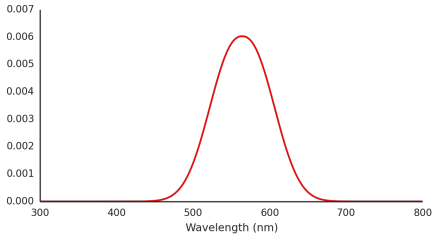
S cone response



X

S cone response

M cone response



X

S cone response

M cone response

L cone response

S cone response

M cone response

L cone response

S cone response

M cone response

L cone response

S cone response

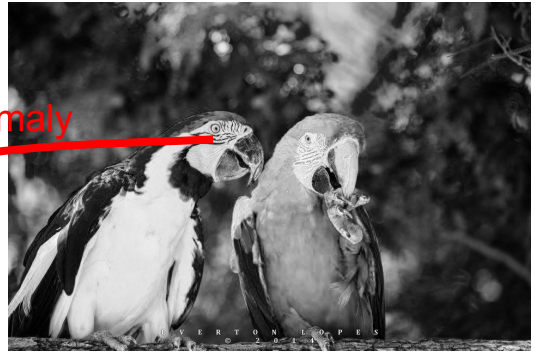
M cone response

L cone response

S cone response

M cone response

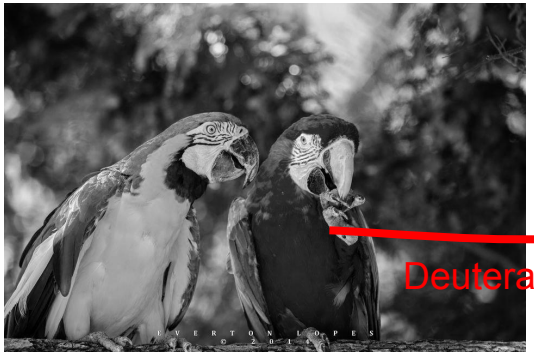
L cone response



Protanomaly

S cone response

M cone response

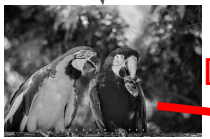


L cone response

Deuteranomaly

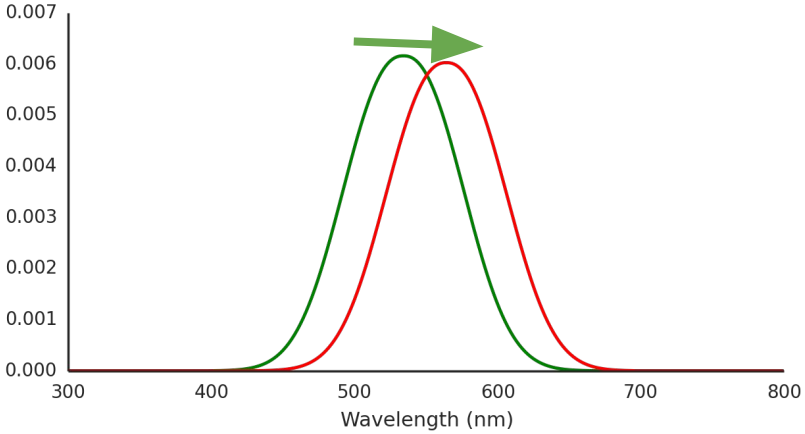
S cone response

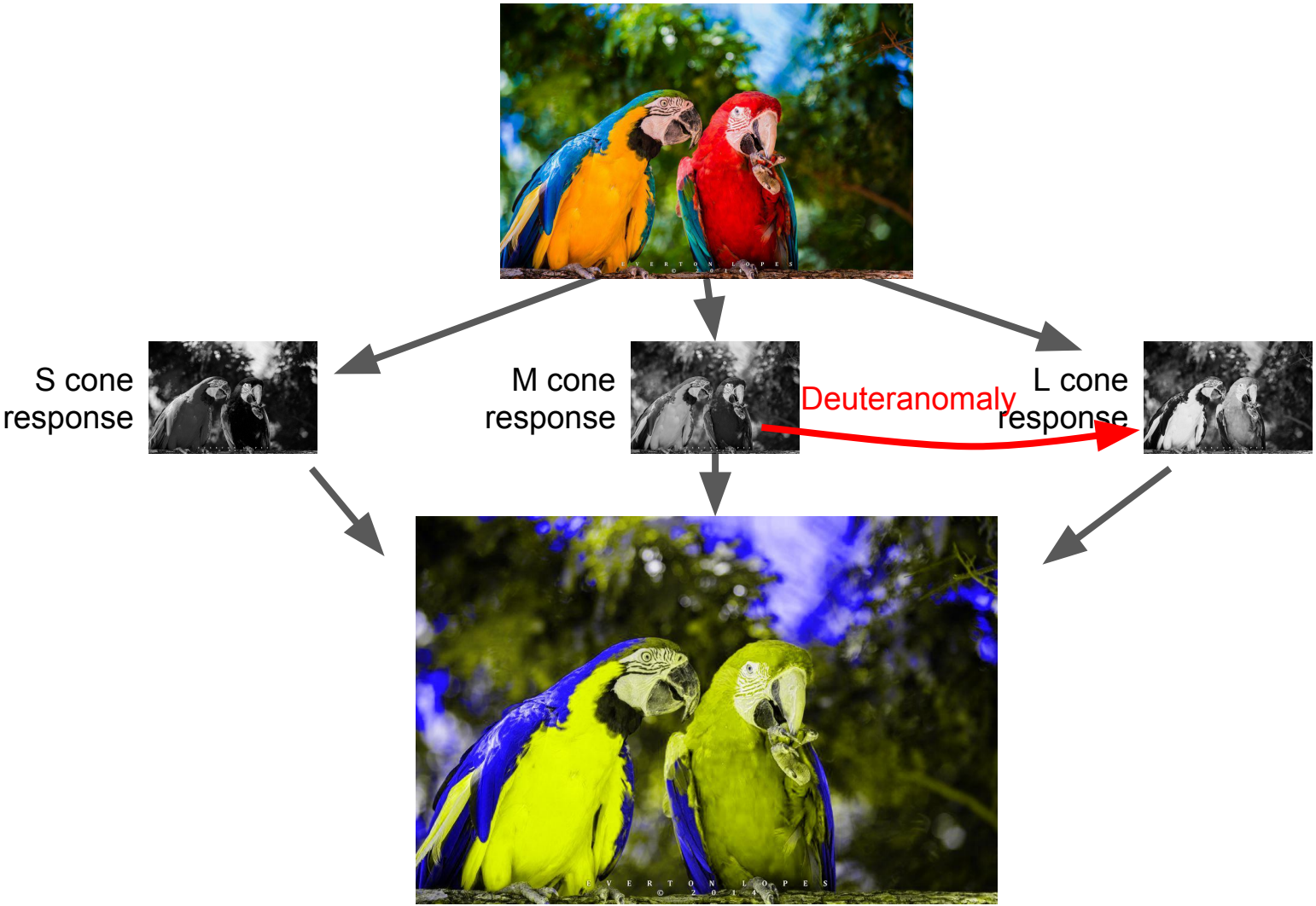
M cone response



Deuteranomaly

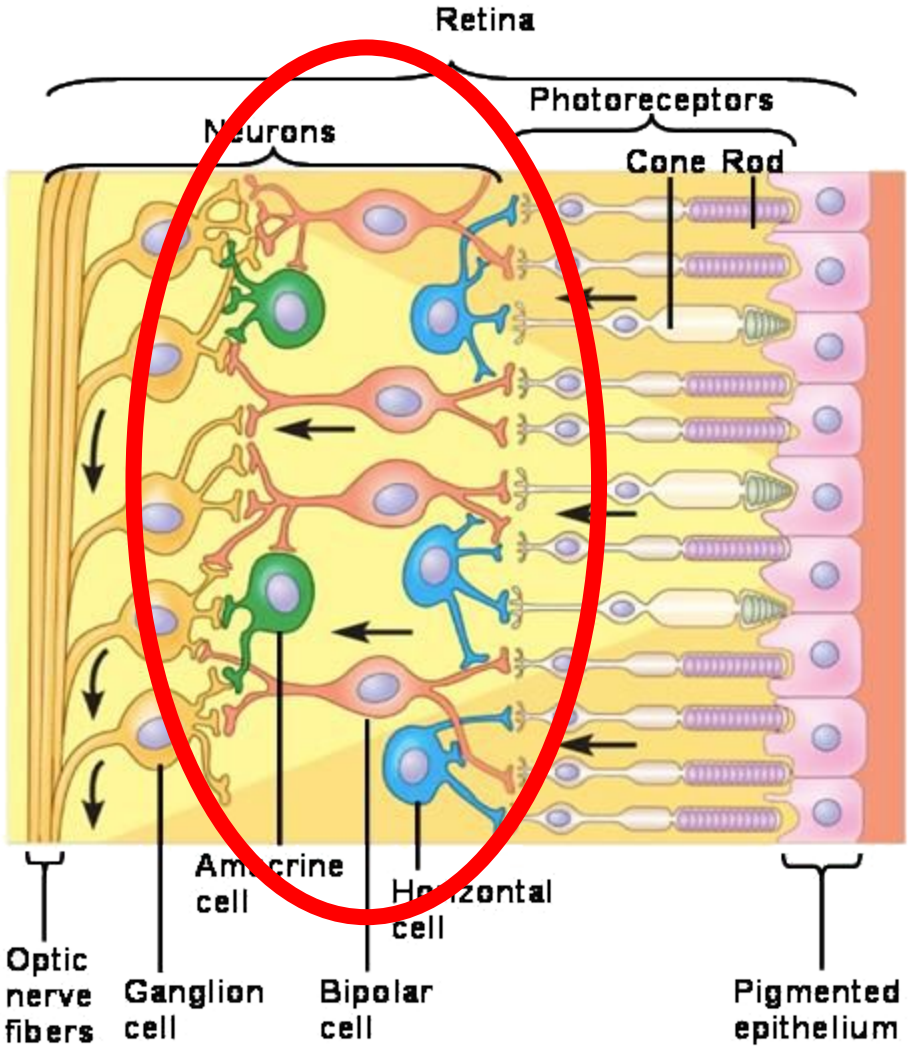
L cone response




S cone response
M cone response
Deuteranomaly
L cone response

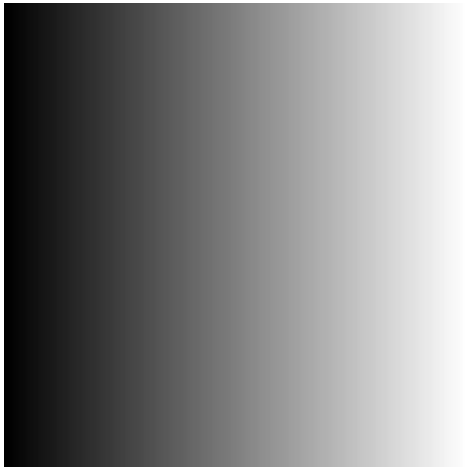
Human Color Vision:

Opponent Process Theory


Retina
Neurons
Photoreceptors
Cone
Rod
Amacrine cell
Horizontal cell
Optic nerve fibers
Ganglion cell
Bipolar cell
Pigmented epithelium

Color Before the Brain

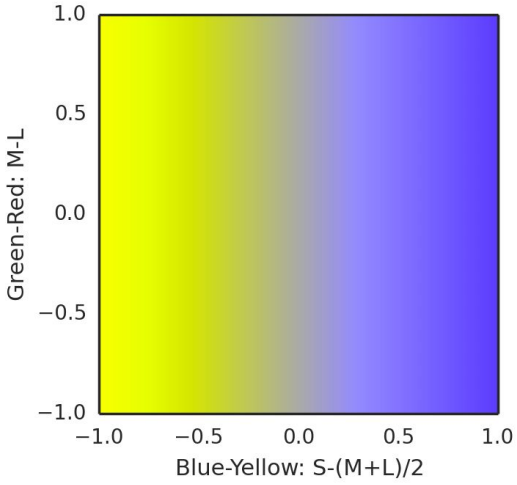
Luminance = S + M + L (+ rods)



Color Before the Brain

Luminance = S + M + L (+ rods)

Blue-Yellow opposition = S - (M + L)

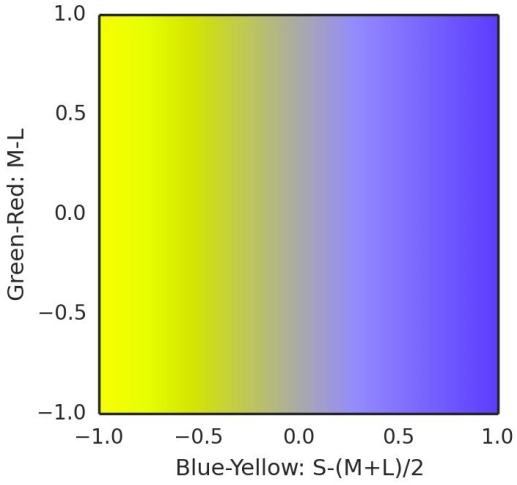


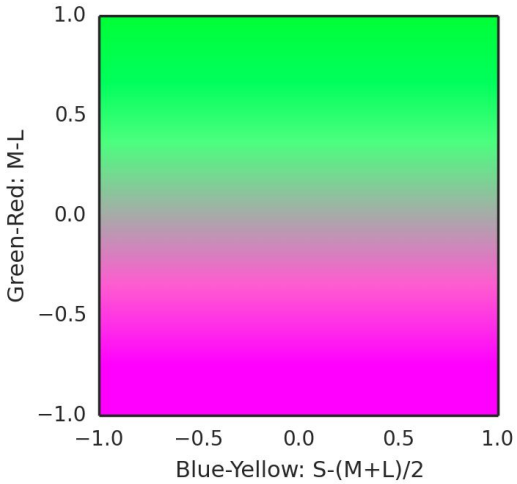
Color Before the Brain

Luminance = S + M + L (+ rods)

Blue-Yellow opposition = S - (M + L)

Red-Green opposition = M - L



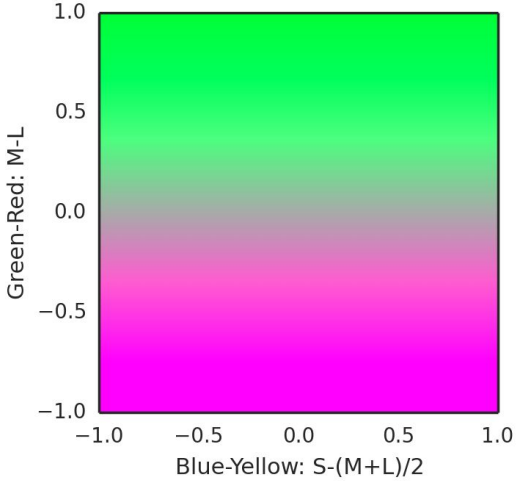


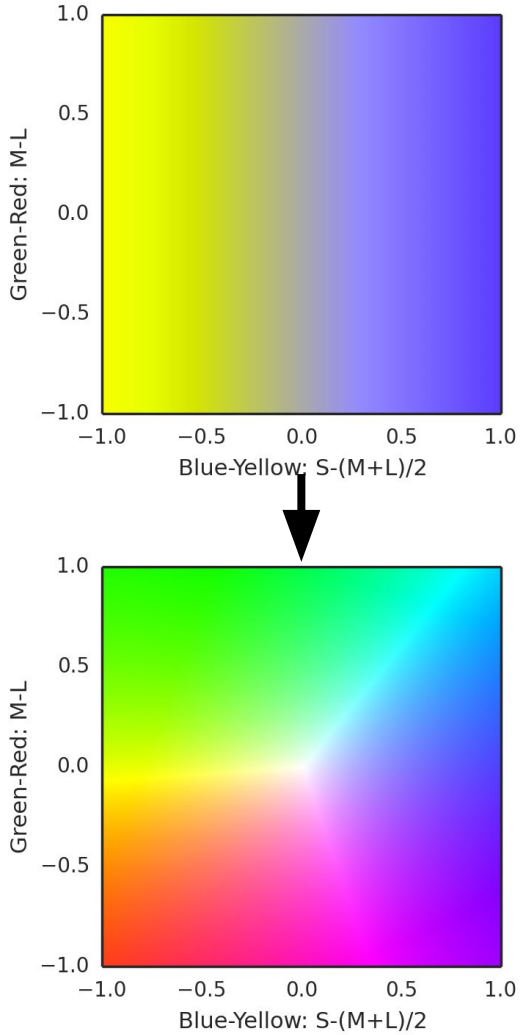
Color Before the Brain

Luminance = S + M + L (+ rods)

Blue-Yellow opposition = S - (M + L)

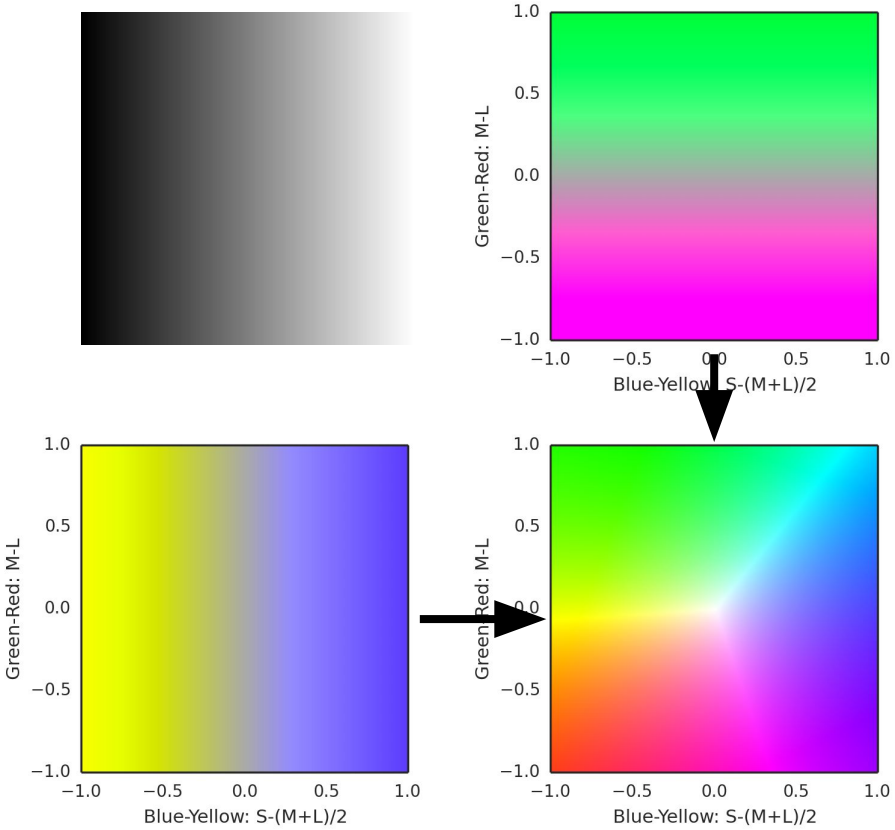
Red-Green opposition = M - L

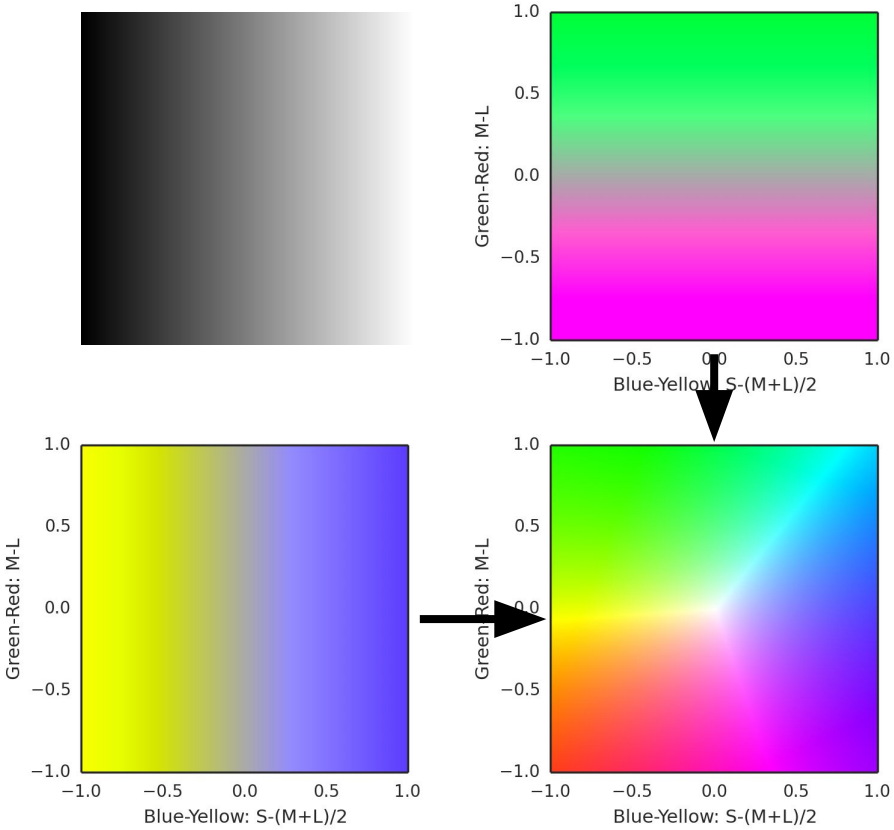


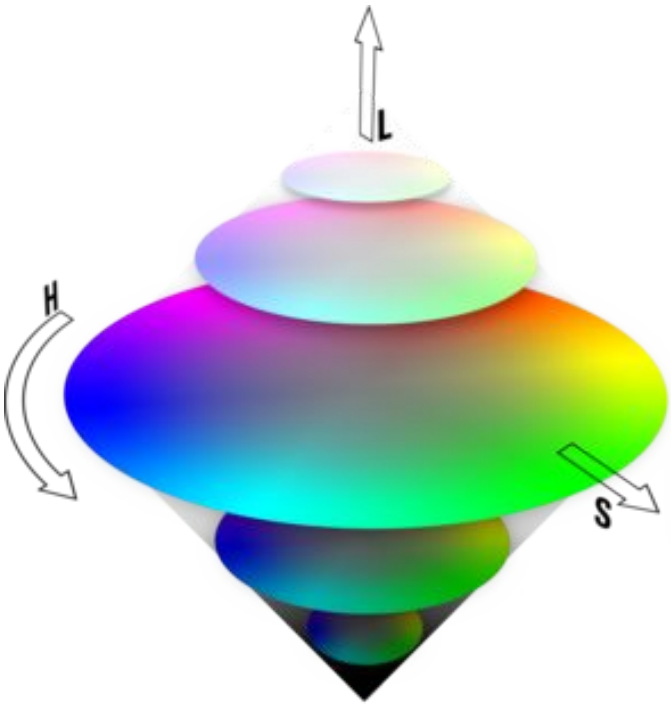


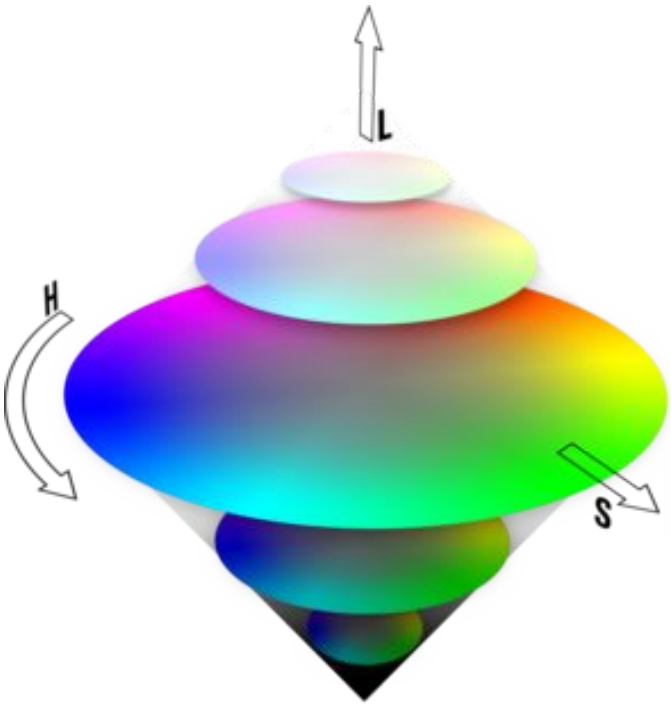
Encoding Using

Hue Saturation and Luminance Space


1.0
0.5
0.0
−0.5
−1.0
Green-Red: M-L
−1.0
−0.5
0.0
0.5
1.0
Blue-Yellow: S-(M+L)/2
Green-Red: M-L
Blue-Yellow: S-(M+L)/2
Green-Red: M-L
Blue-Yellow: S-(M+L)/2


1.0
0.5
0.0
−0.5
−1.0
Green-Red: M-L
−1.0
−0.5
0.0
0.5
1.0
Blue-Yellow: S-(M+L)/2


L
H
S

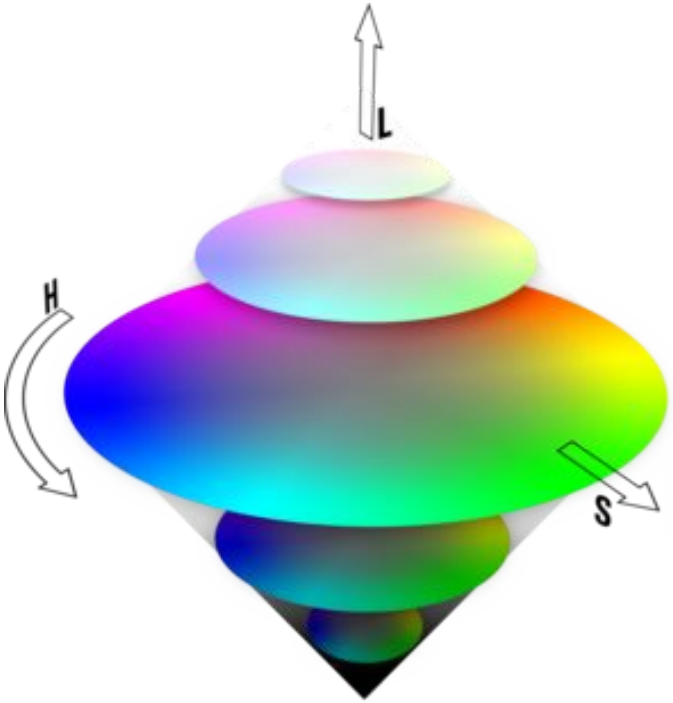

L
H
S

HSL Encoding

Hue

Saturation (and opposing processes)

Luminance/Value



HSL Encoding

Hue:

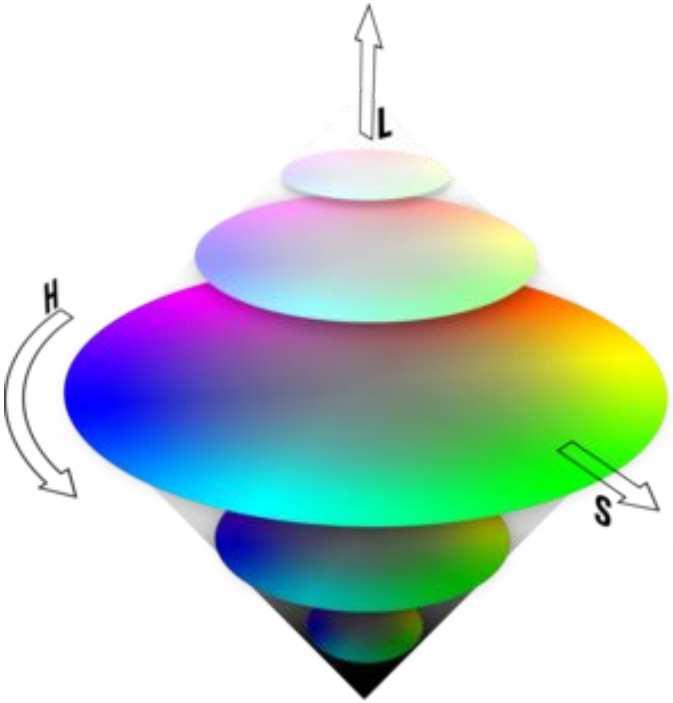
Nominal

Saturation (and opposing processes):

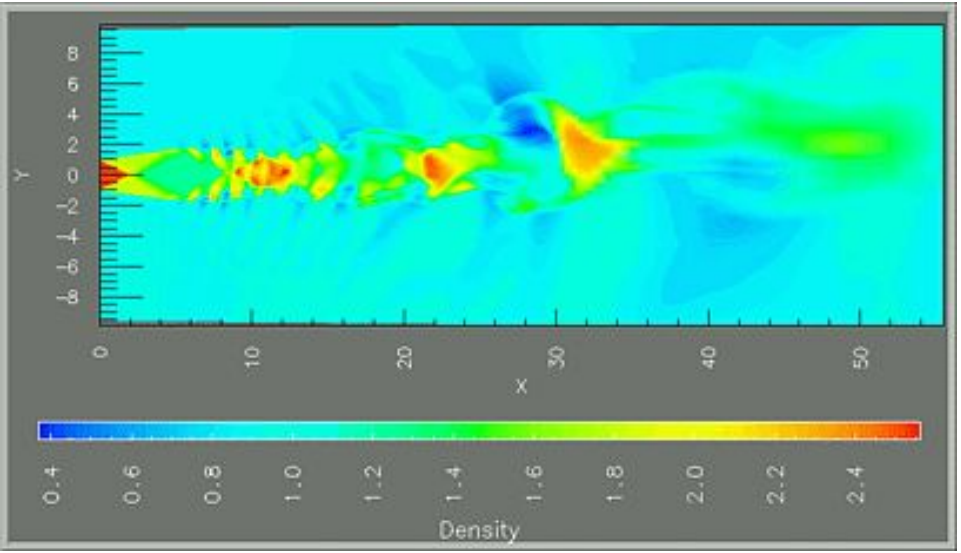
Quantitative

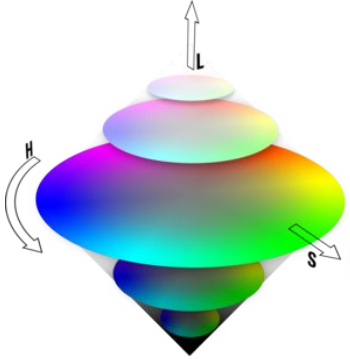
Luminance/Value:

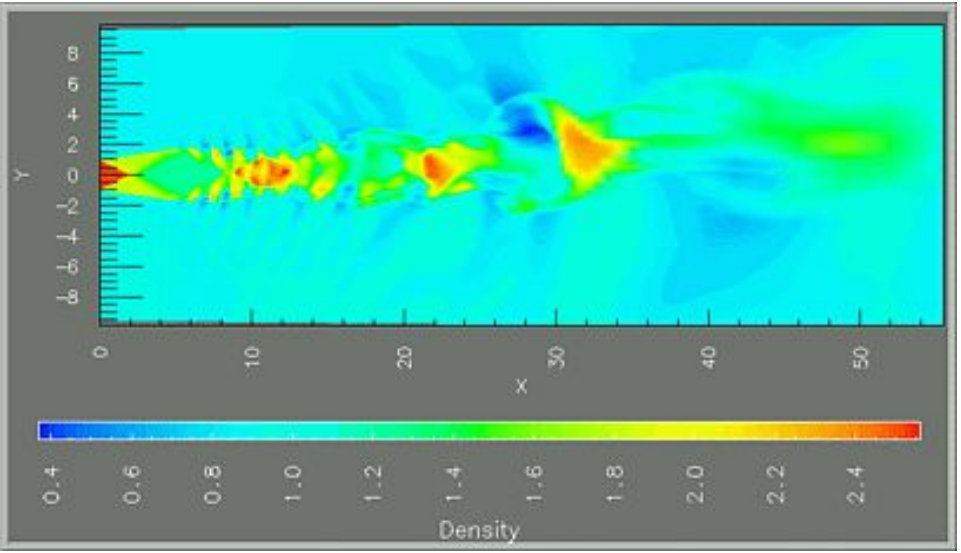
Quantitative

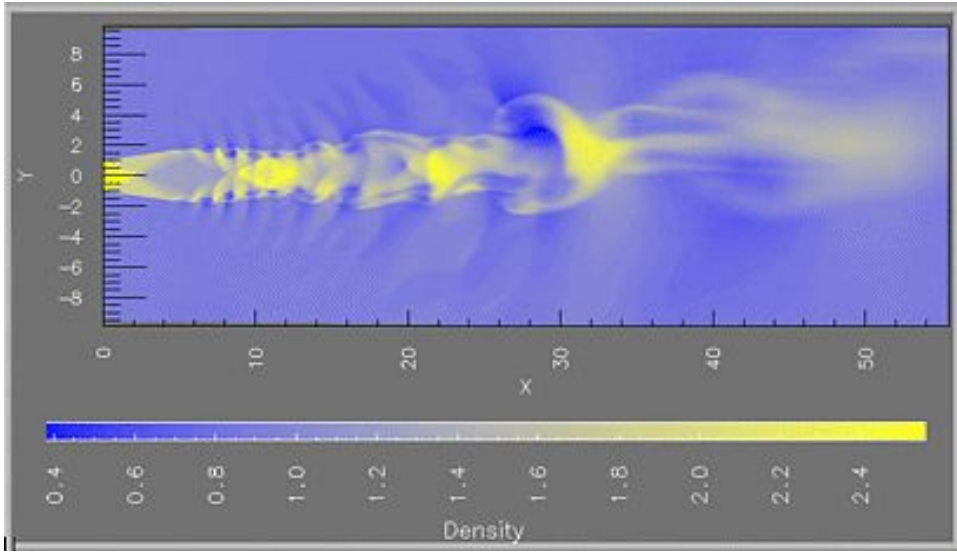


Examples

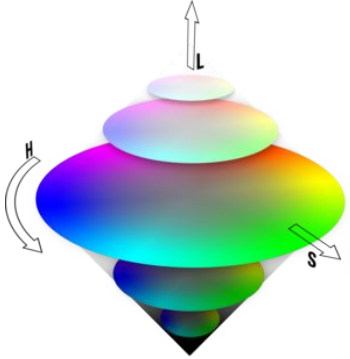

Y
8
6
4
2
0
-2
-4
-6
-8
0
10
20
30
40
50
X
0.4
0.6
0.8
1.0
1.2
1.4
1.6
1.8
2.0
2.2
2.4
Density

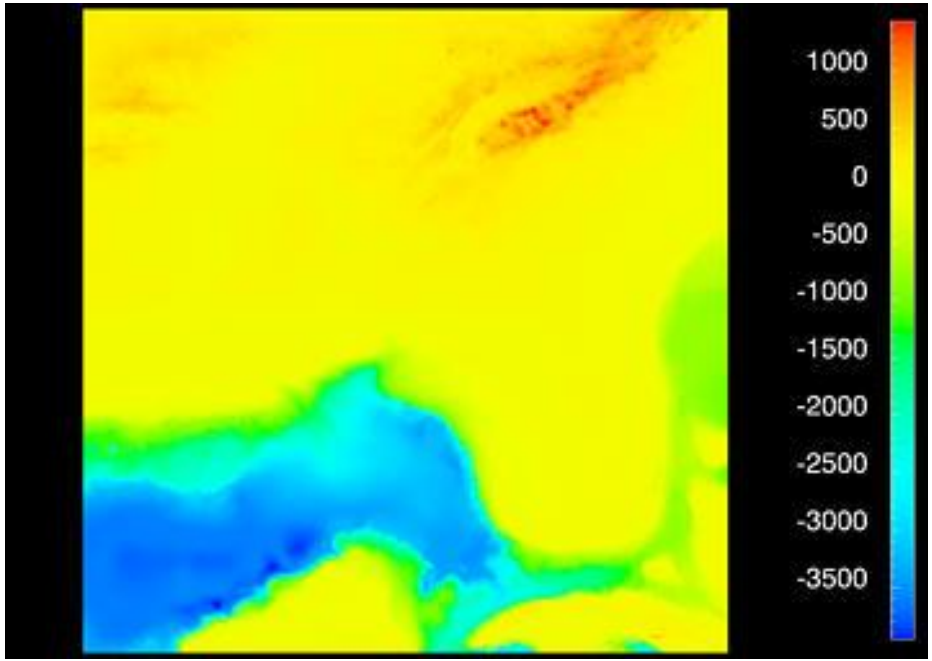

L
H
S


Y
8
6
4
2
0
-2
-4
-6
-8
0
10
20
30
40
50
X
0.4
0.6
0.8
1.0
1.2
1.4
1.6
1.8
2.0
2.2
2.4
Density

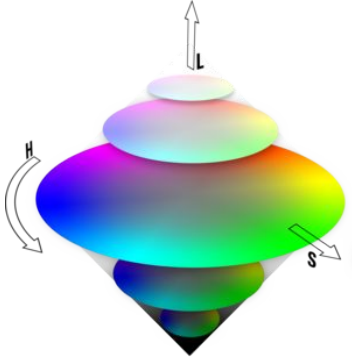

Y
8
6
4
2
0
-2
-4
-6
-8
0
10
20
30
40
50
X
0.4
0.6
0.8
1.0
1.2
1.4
1.6
1.8
2.0
2.2
2.4
Density

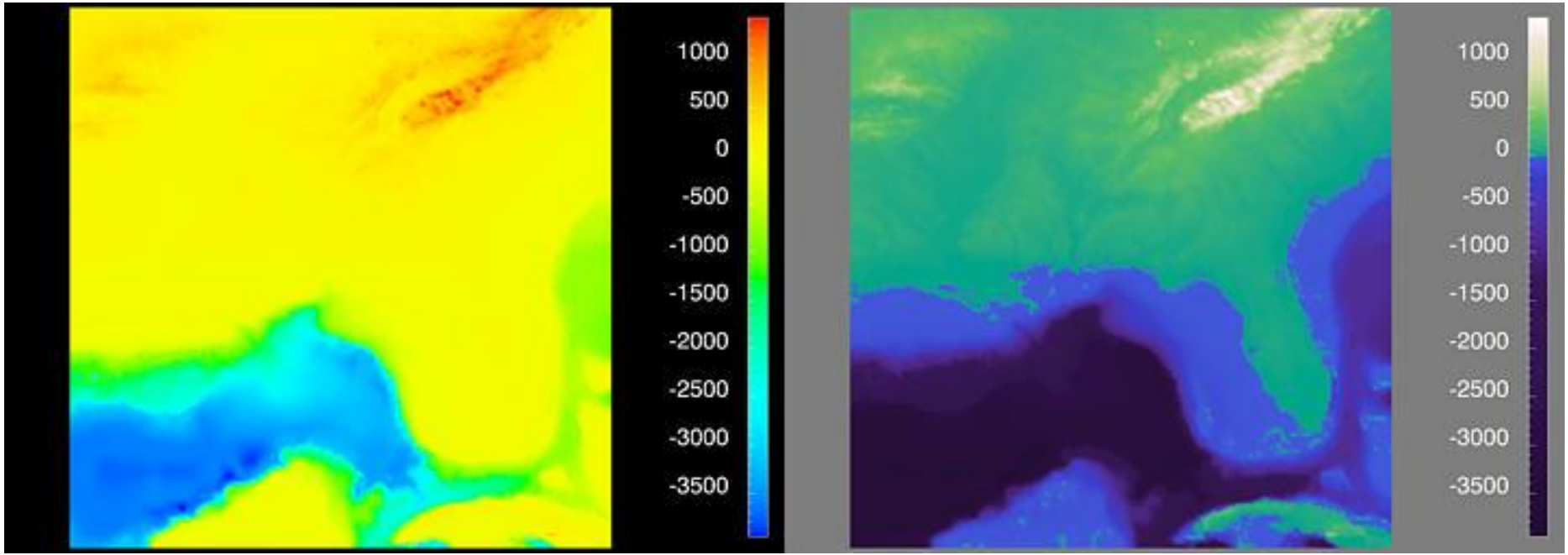
Elevation Example



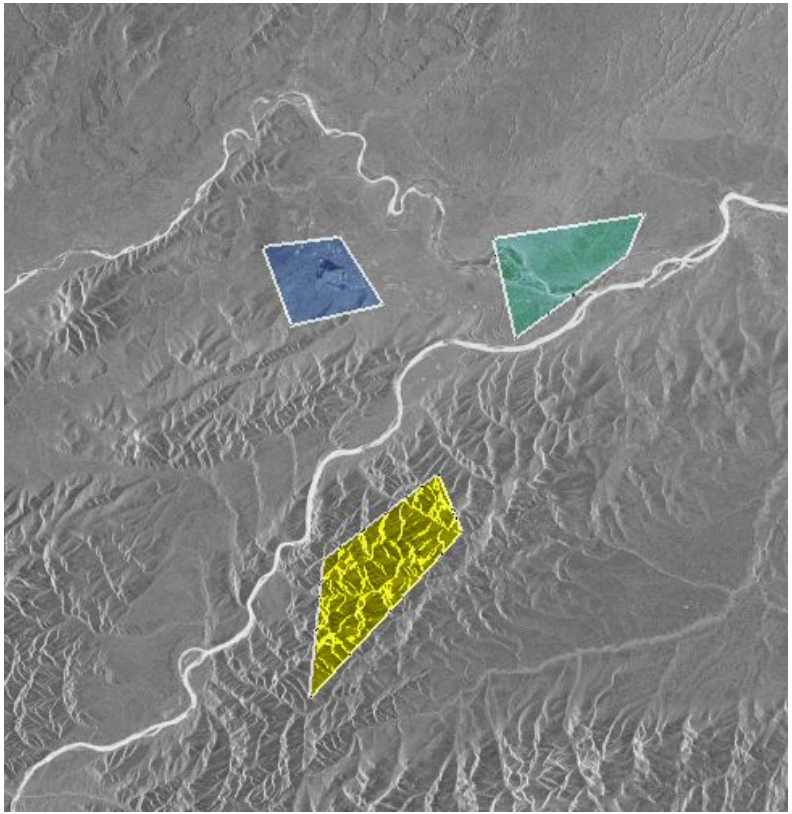


Elevation Example





Example: Encoding multiple dimensions



HSL Encoding: Which Quantitative Where?

Hue:

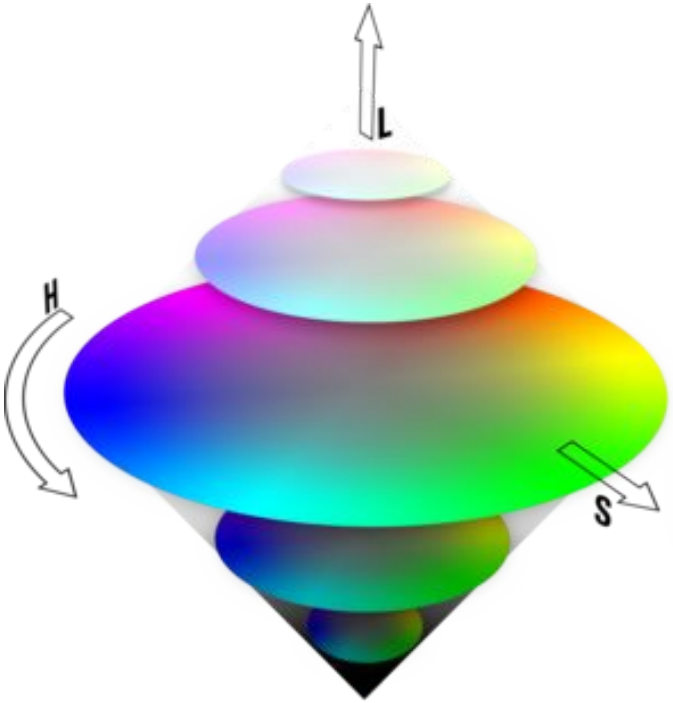
Nominal

Saturation (and opposing processes):

Quantitative

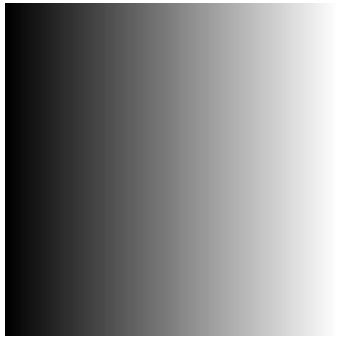
Luminance/Value:

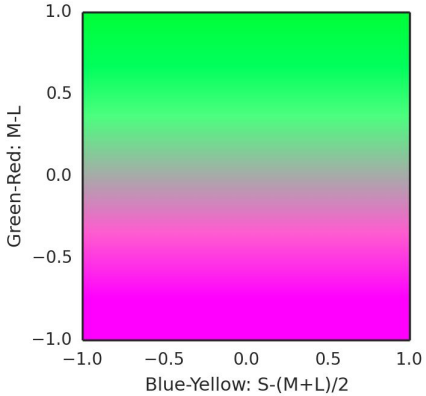
Quantitative

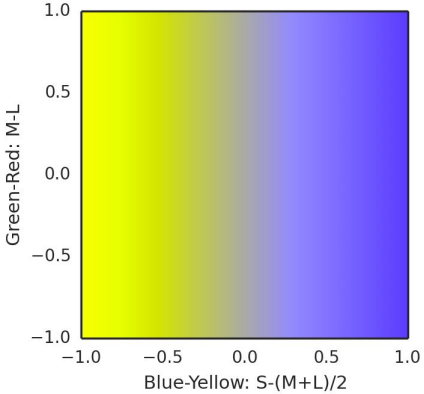


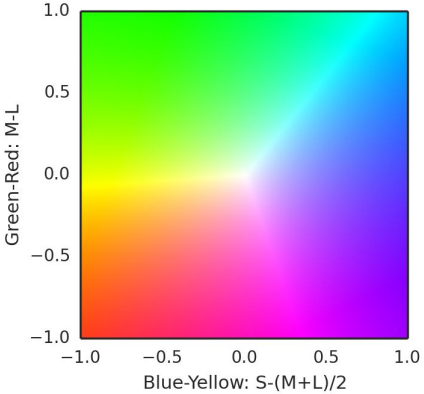
High vs Low Frequency Data

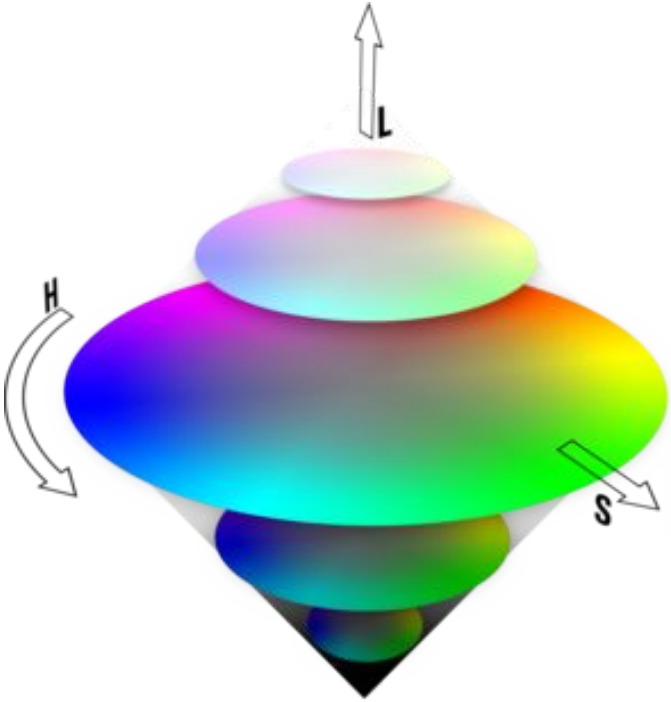
HSV/HSL: Three Different Dimensions











Original

+

Red-Green

+

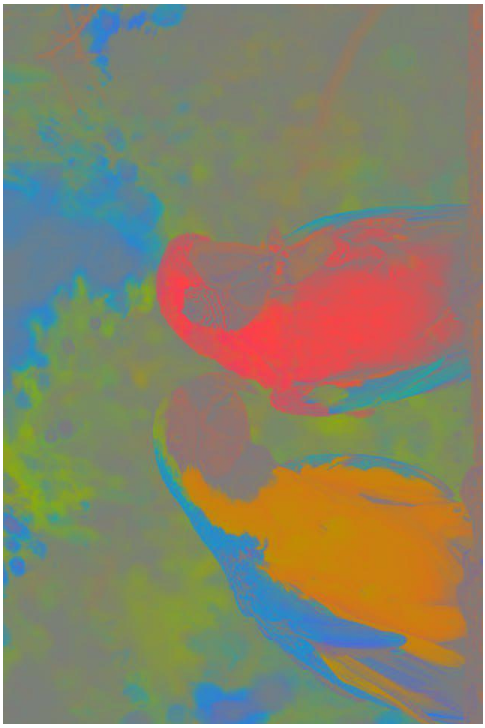
Luminance

=

Blue-Yellow

=

Chrominance
(Blue-Yellow and Red-Green)



Original

+

Luminance

=

Luminance

+

Chrominance
(Blue-Yellow and Red-Green)

=

Compressed

66%

≈

2x downsample

Original

Original

=

Chrominance
(Blue-Yellow and Red-Green)

4x downsample

+

Luminance

≈

Compressed

50%

Original

=

Chrominance
(Blue-Yellow and Red-Green)

8x downsample

+

Luminance

≈

Compressed

42%

Original

=

Chrominance
(Blue-Yellow and Red-Green)

16x downsample

+

Luminance

≈

Compressed

38%

Original

=

Chrominance
(Blue-Yellow and Red-Green)

32x downsample

+

Luminance

≈

Compressed

35%

Original

≈

Compressed

34%

=

Chrominance
(Blue-Yellow and Red-Green)

64x downsample

+

Luminance

HSL Encoding

Hue:

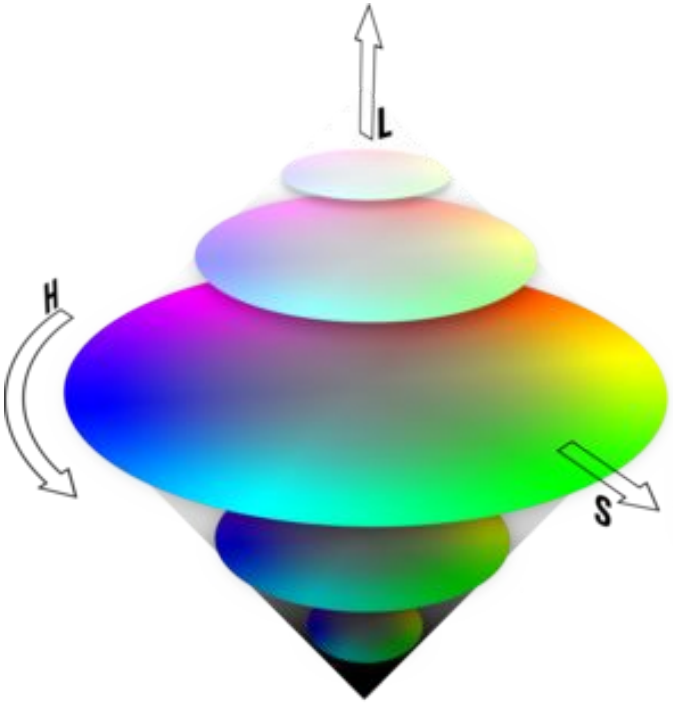
Nominal

Saturation (and opposing processes):

Low Frequency Quantitative

Luminance/Value:

High Frequency Quantitative



Perceptual Uniformity

Weber’s law (adulterated): The just-noticeable difference (JND) between two stimuli is proportional to the magnitude of the stimuli.

Perceptually uniform: A change of the same amount in a color value should produce a change of about the same visual importance.

Perceptual Uniformity

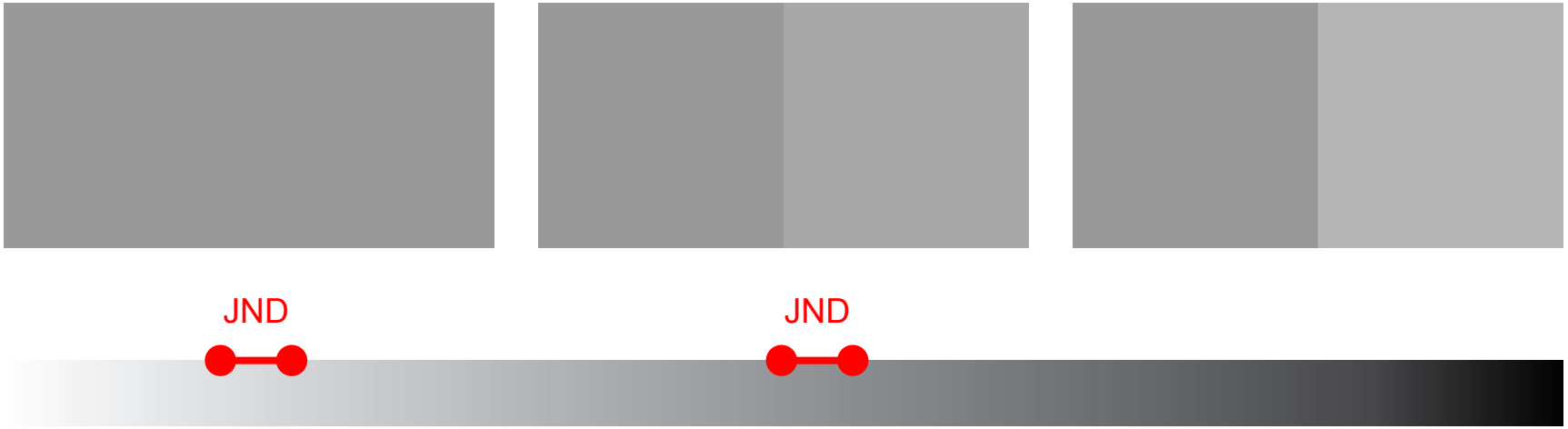
Weber’s law (adulterated): The just-noticeable difference (JND) between two stimuli is proportional to the magnitude of the stimuli.

Perceptually uniform: A change of the same amount in a color value should produce a change of about the same visual importance.

Perceptual Uniformity

Weber’s law (adulterated): The just-noticeable difference (JND) between two stimuli is proportional to the magnitude of the stimuli.

Perceptually uniform: A change of the same amount in a color value should produce a change of about the same visual importance.



Colormaps in Matplotlib

Considerations:

- Colorblindness
- Perceptual uniformity
- Print compatibility

http://matplotlib.org/examples/color/colormaps_reference.html

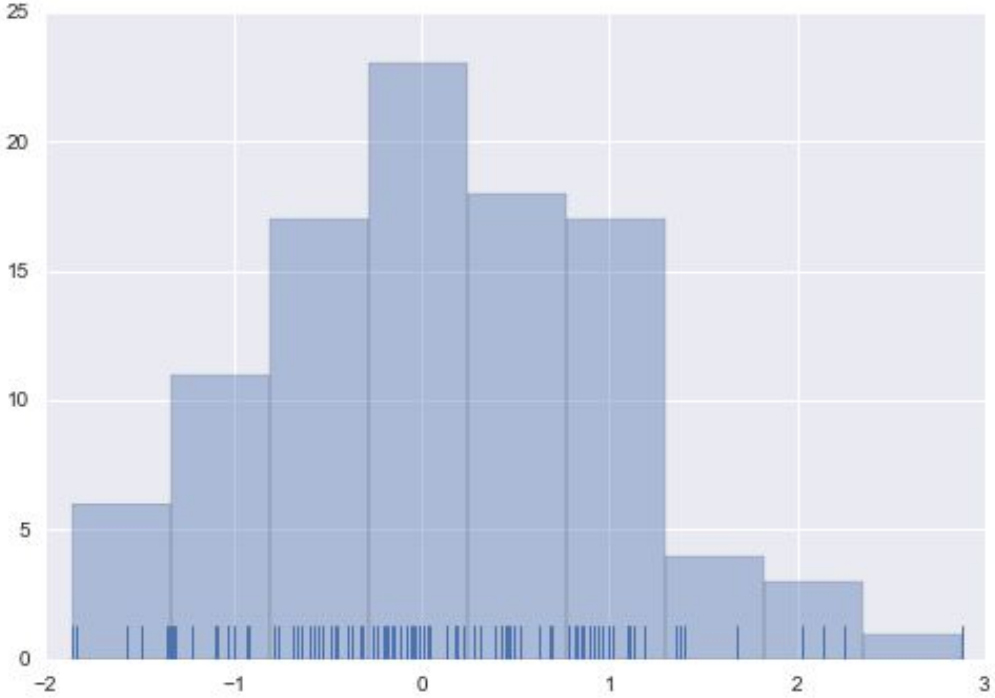
http://matplotlib.org/users/dflt_style_changes.html

Color Takeaways

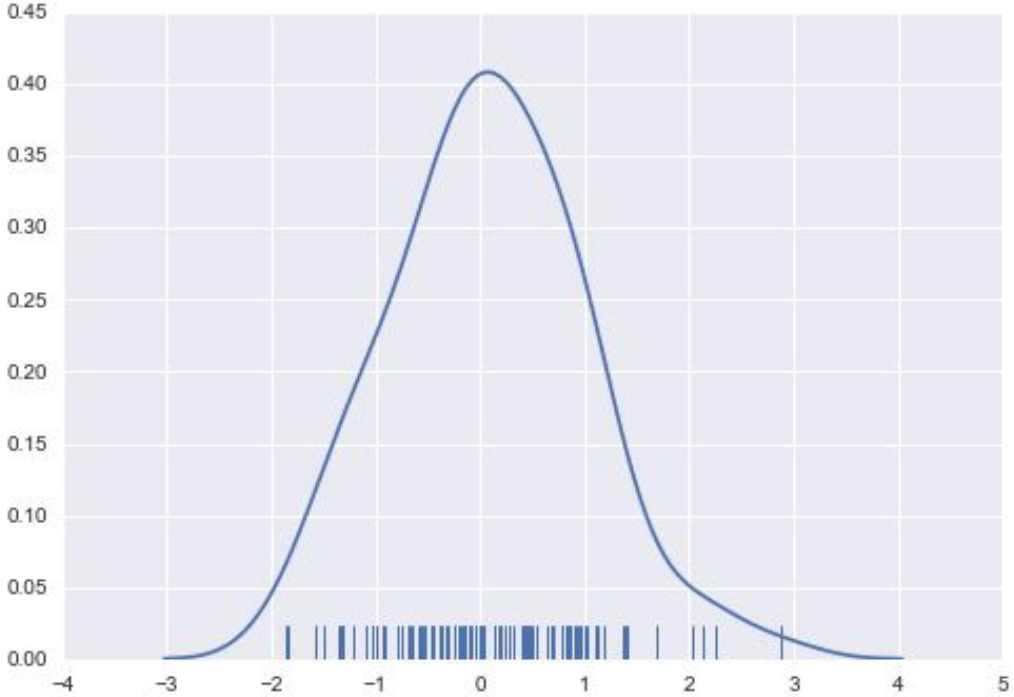
- Never use the jet colormap!
- Use a colormap that is
  - perceptually uniform
  - colorblindness-insensitive
- When making your own color maps:
  - Hue: Nominal
  - Saturation: Low frequency quantitative
  - Luminance: High frequency quantitative
  - Segmentation for context

Visualizing Distributions

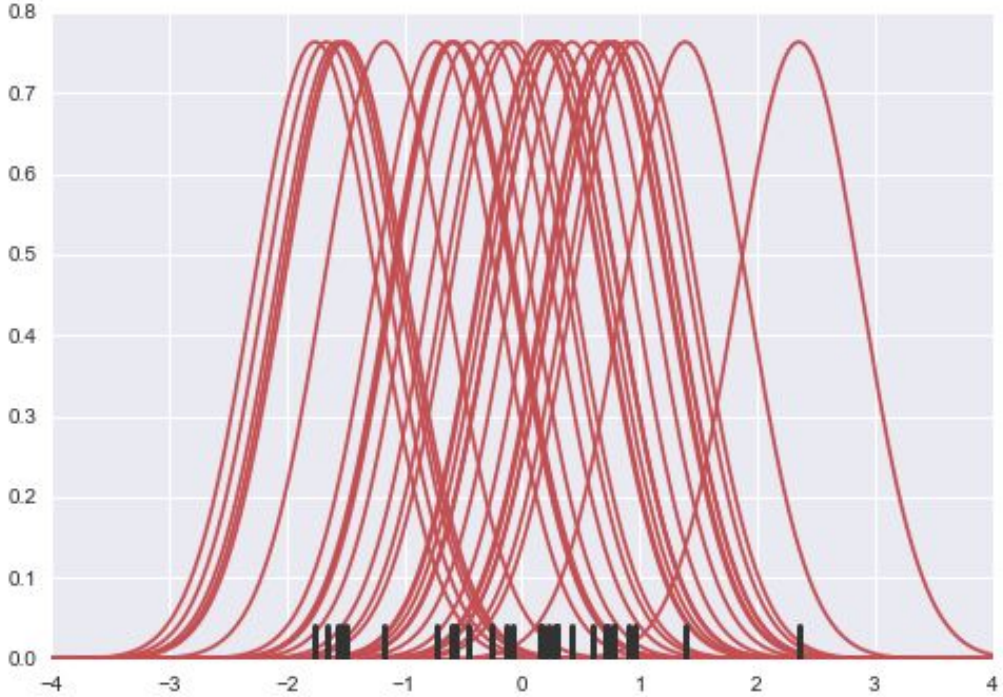
Histograms



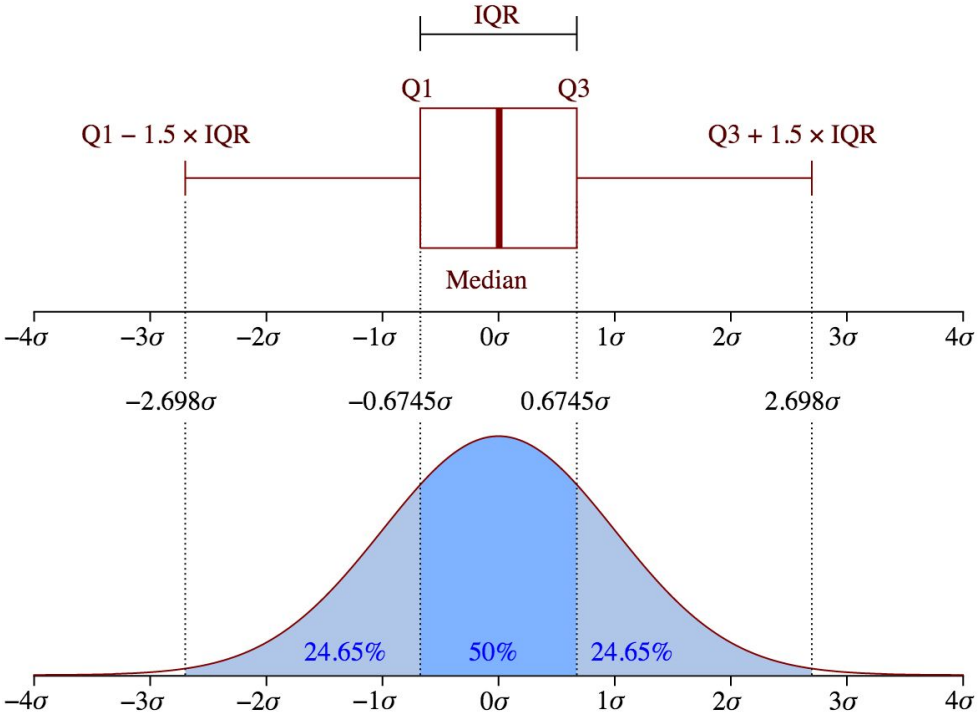
Kernel Density Estimate Plots



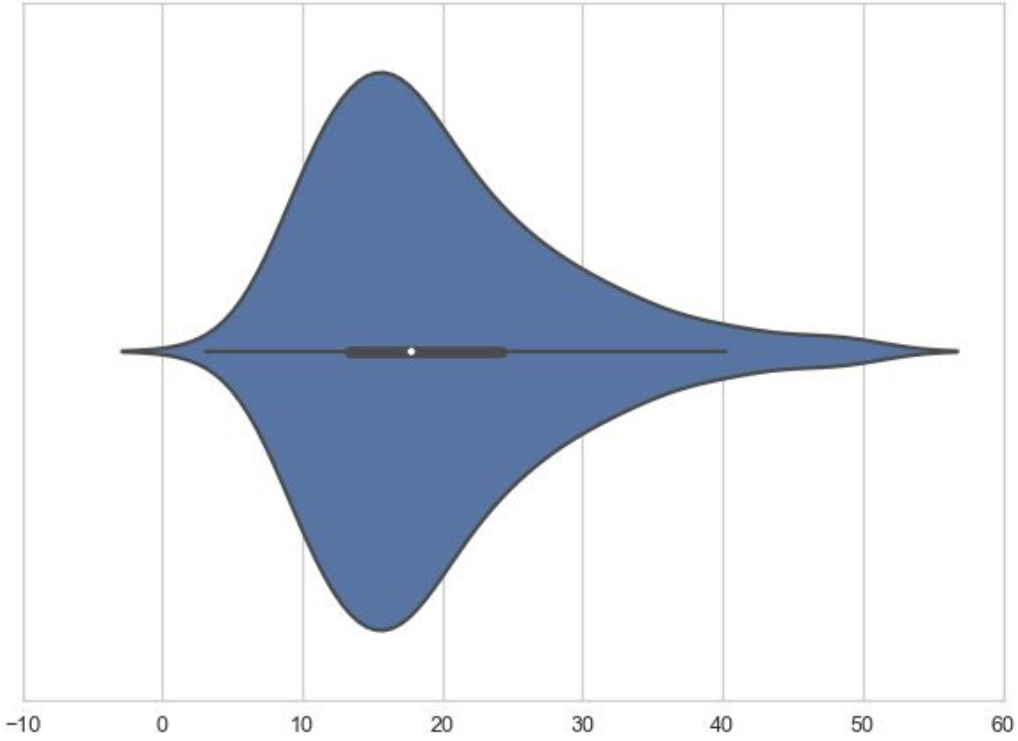
Kernel Density Estimate Plots



Box Plots


IQR
Q1
Q3
Q1 − 1.5 × IQR
Q3 + 1.5 × IQR
Median
−4σ
−3σ
−2σ
−1σ
0σ
1σ
2σ
3σ
4σ
−2.698σ
−0.6745σ
0.6745σ
2.698σ
24.65%
50%
24.65%

Violin Plots



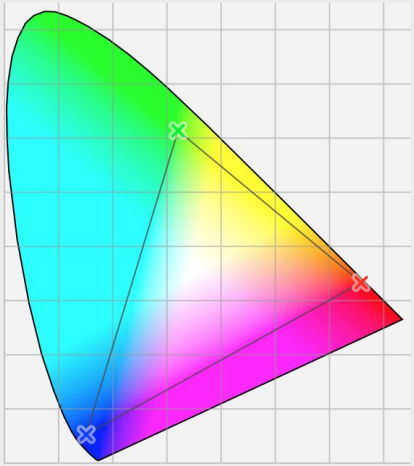
Notebook Example

Visualizing High Dimensional Data

3D Visualization Gone Wrong

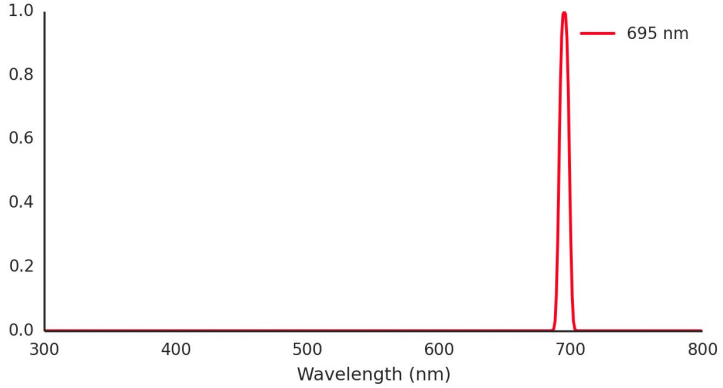
Color LCD-C6076DC3-B2A7-E5A5-3A31-D9E7427CFF58.icc

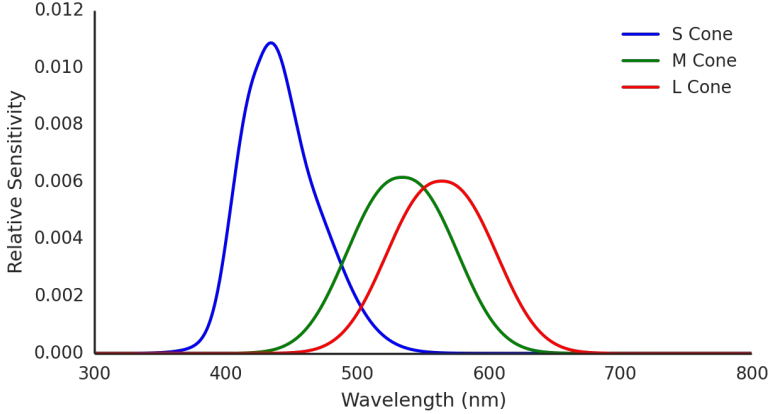
#	Tag	Data	Size	Description
4	'wtpt'	'XYZ '	20	Media white-point tristimulus
5	'rXYZ'	'XYZ '	20	Red colorant tristimulus
6	'gXYZ'	'XYZ '	20	Green colorant tristimulus
7	'bXYZ'	'XYZ '	20	Blue colorant tristimulus
8	'rTRC'	'curv'	2,060	Red tone response curve
9	'bTRC'	'curv'	2,060	Blue tone response curve
10	'gTRC'	'curv'	2,060	Green tone response curve

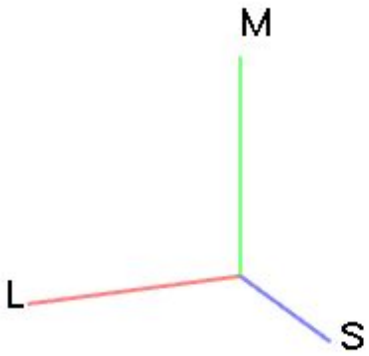


Red	X:	0.444
	Y:	0.225
	Z:	0.005
Green	X:	0.379
	Y:	0.726
	Z:	0.078
Blue	X:	0.140
	Y:	0.049
	Z:	0.741

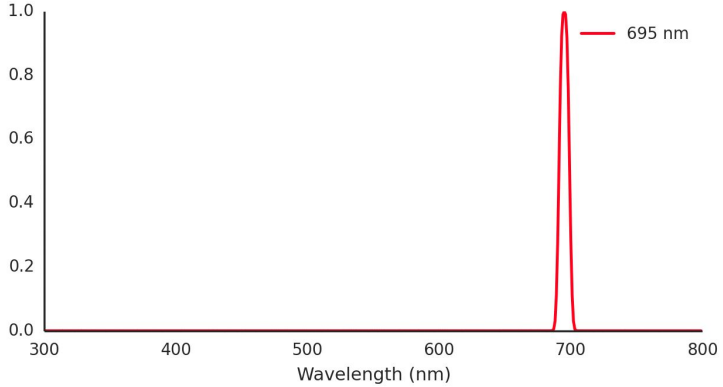
Spectral Color Tristimulus Values

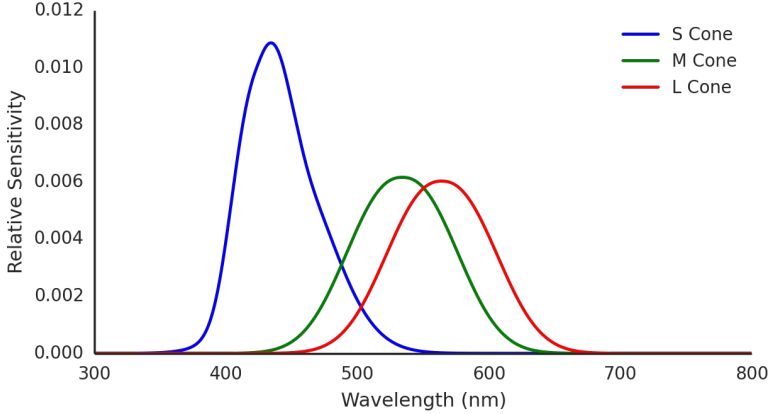


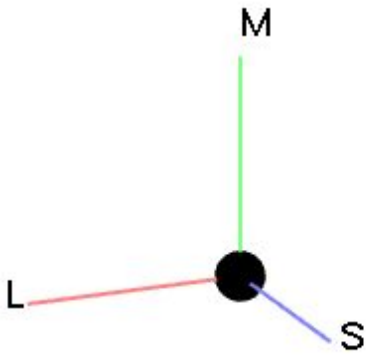




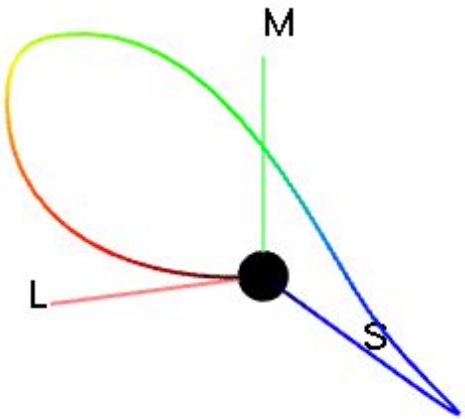
Spectral Color Tristimulus Values



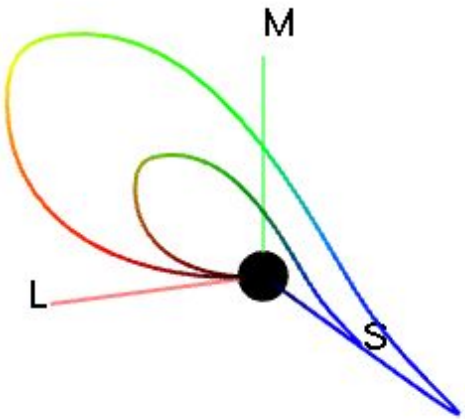




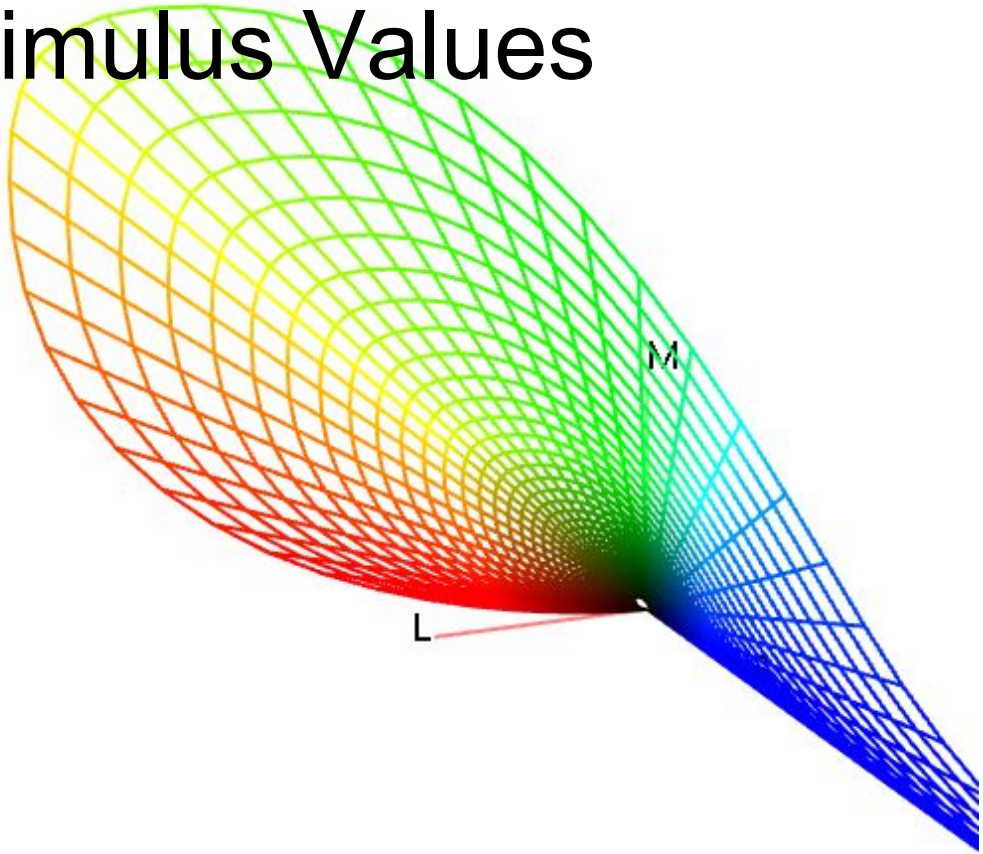
Spectral Color Tristimulus Values

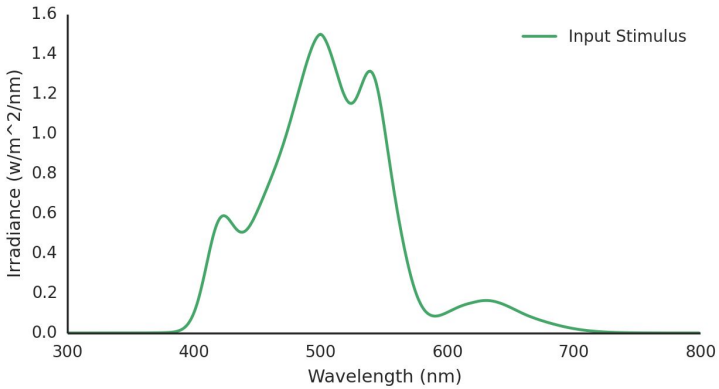


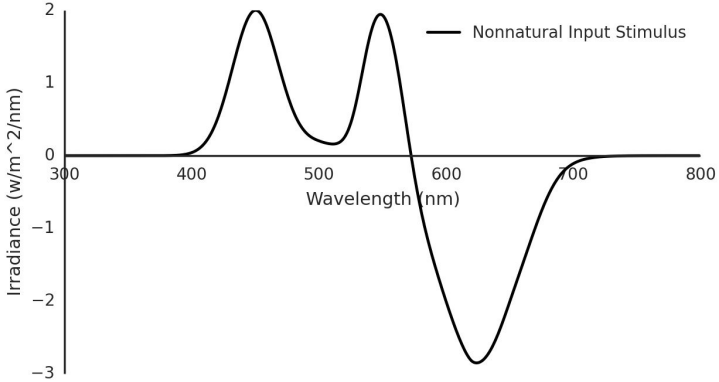
Spectral Color Tristimulus Values

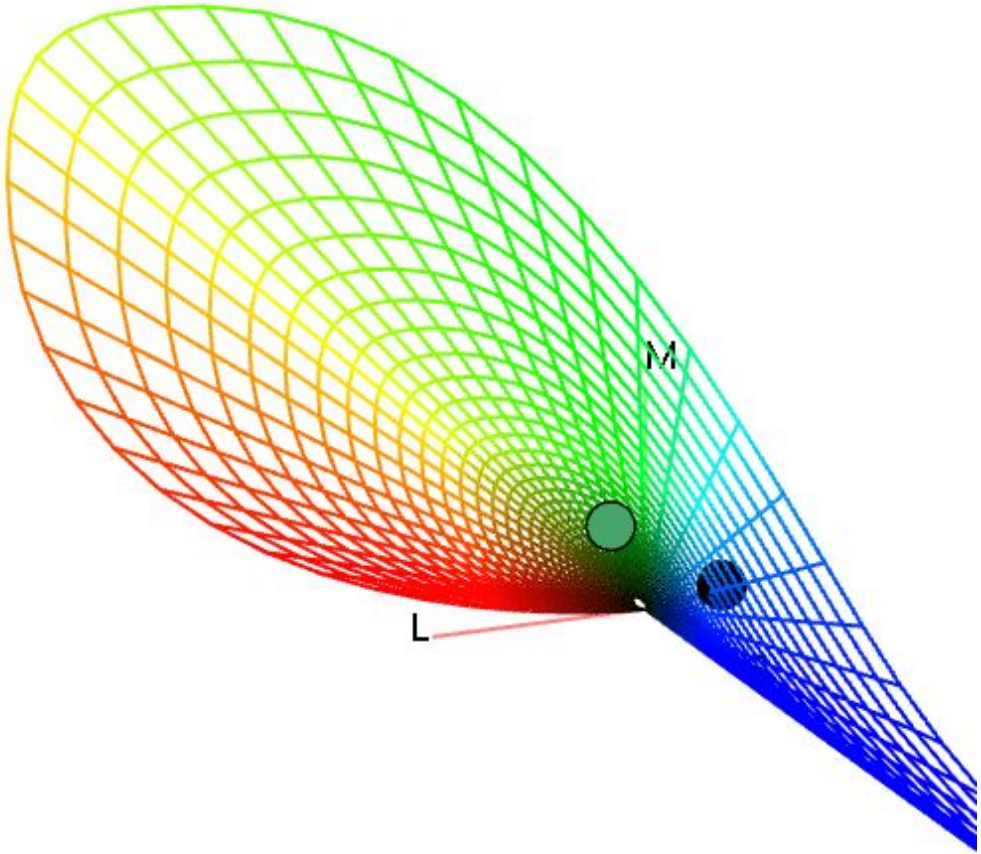


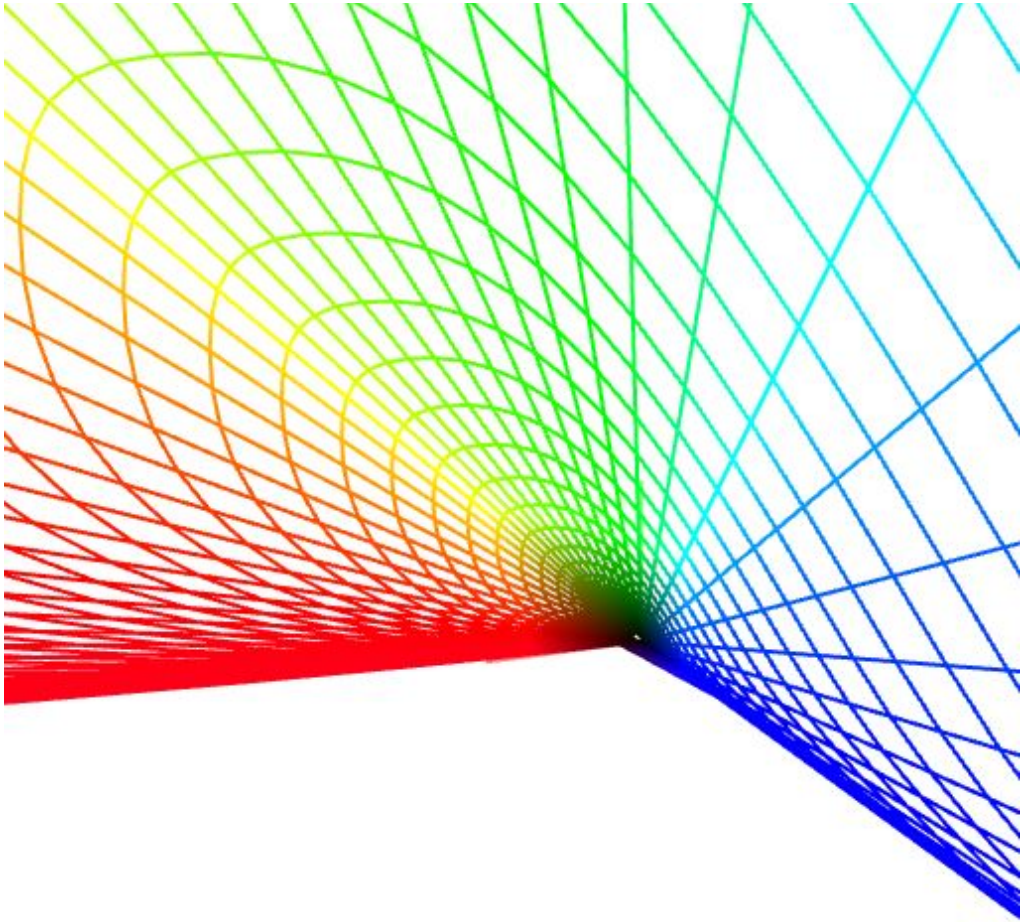
Spectral Color Tristimulus Values

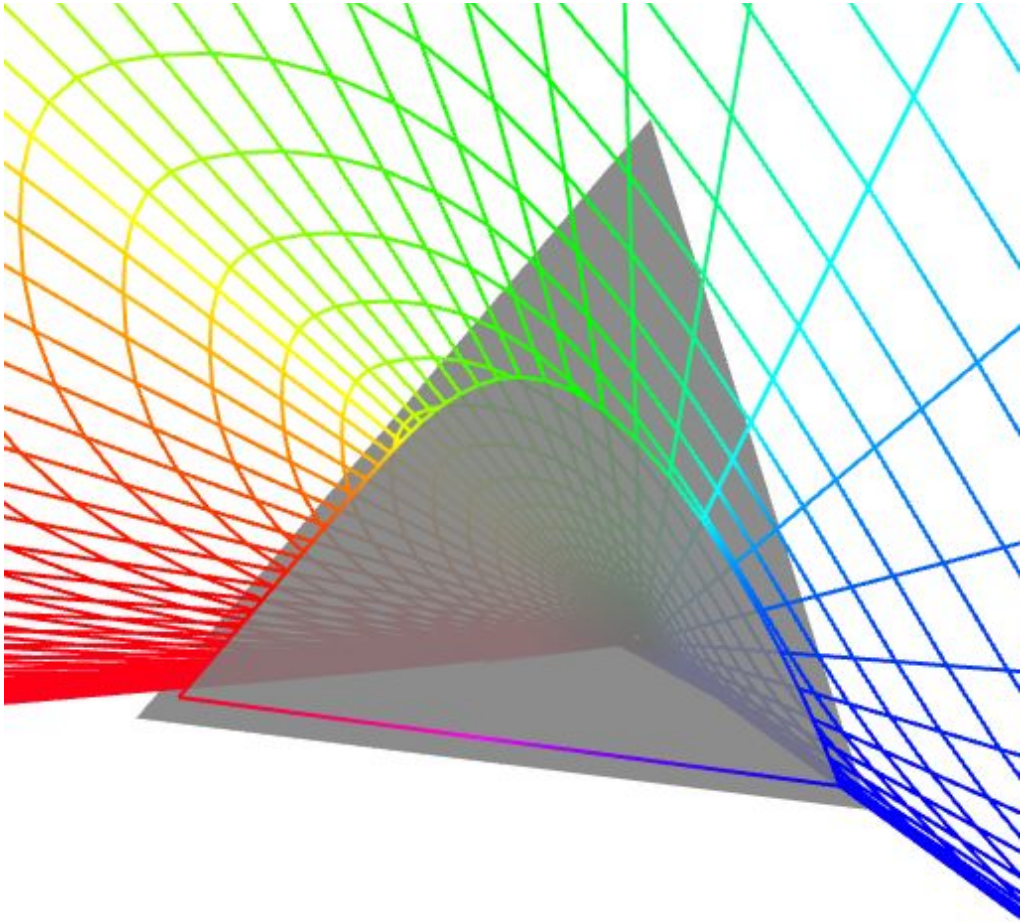


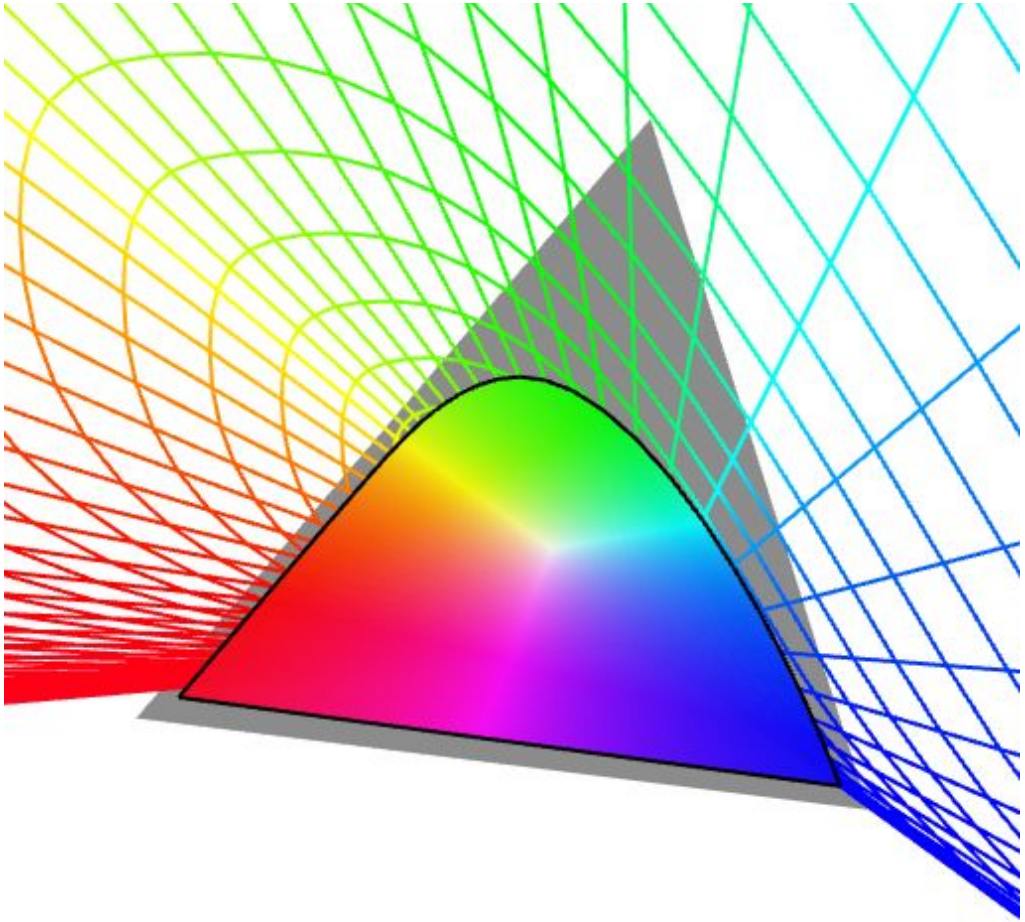

Input Stimulus
Irradiance (w/m^2/nm)
Wavelength (nm)
1.6
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
300
400
500
600
700
800

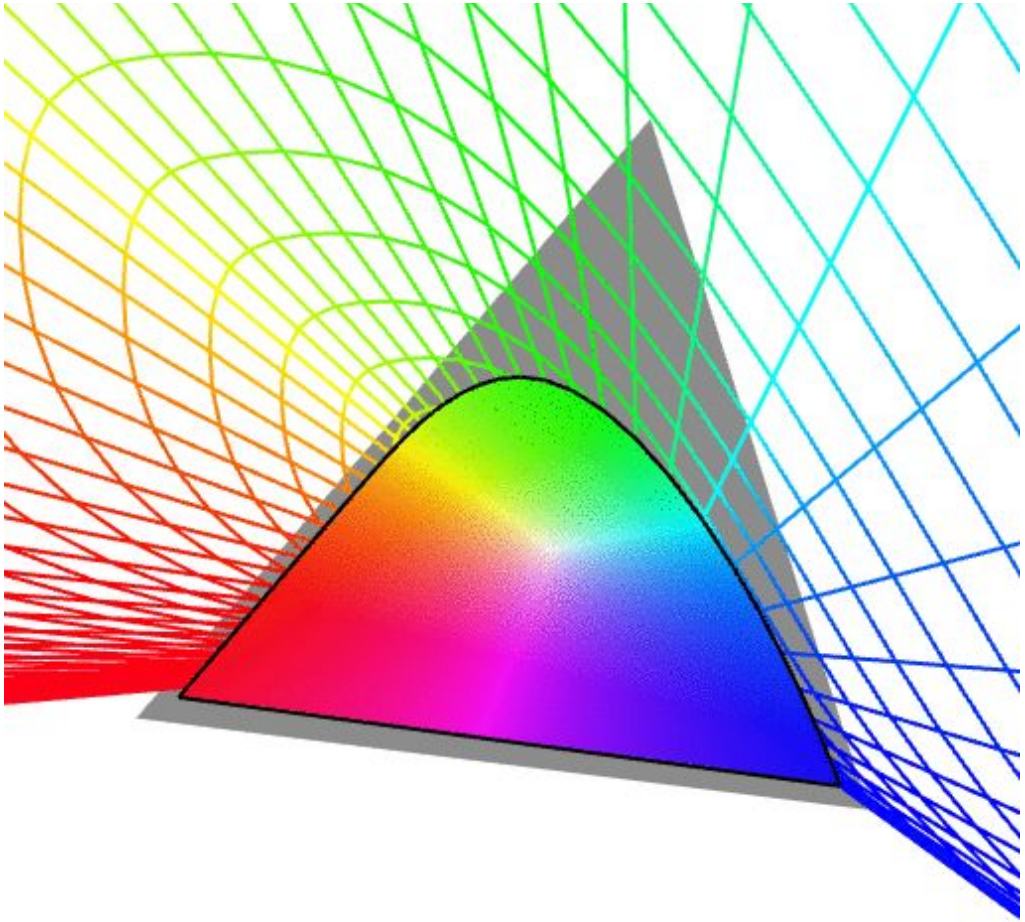

Nonnatural Input Stimulus
Irradiance (w/m^2/nm)
Wavelength (nm)
2
1
0
−1
−2
−3
300
400
500
600
700
800

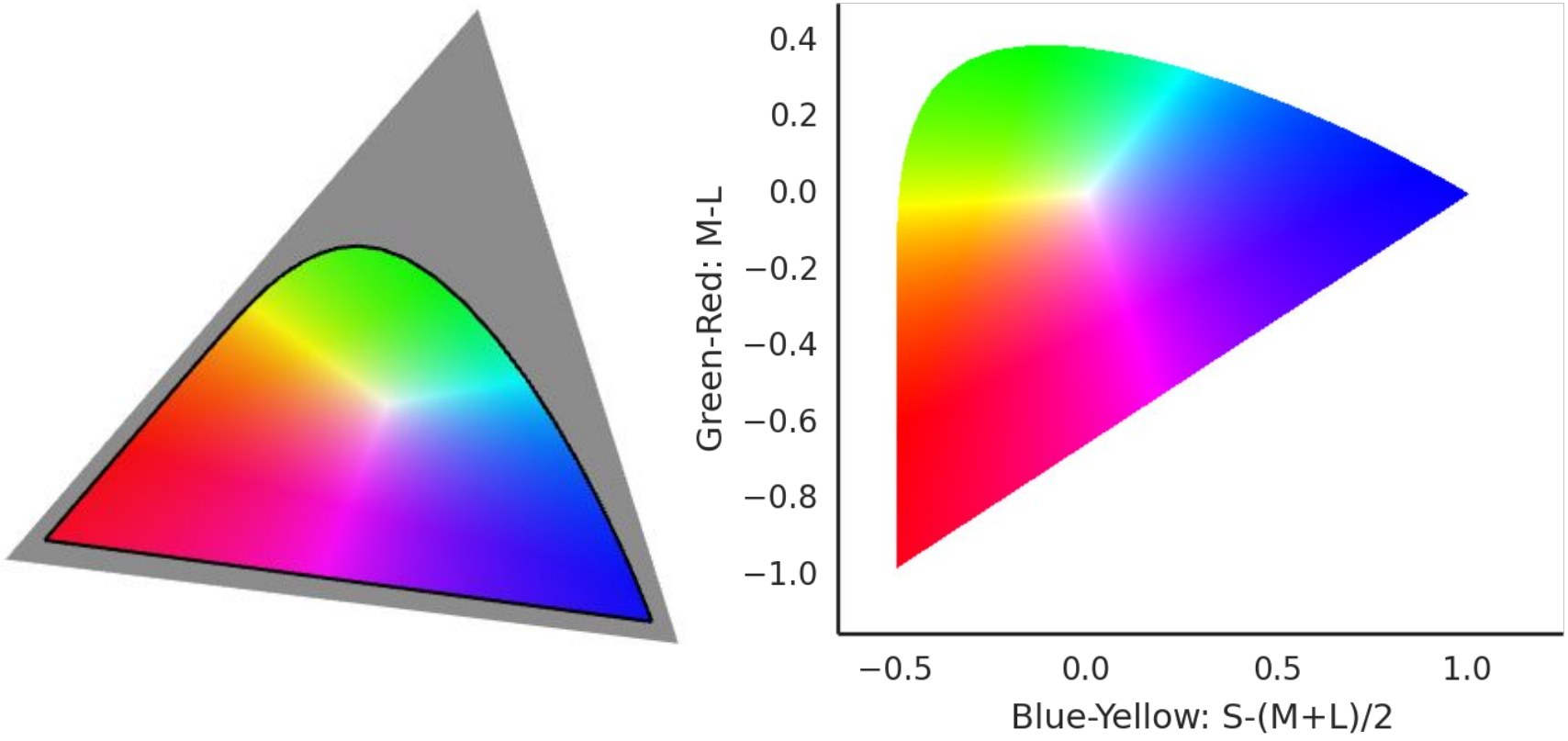

M
L

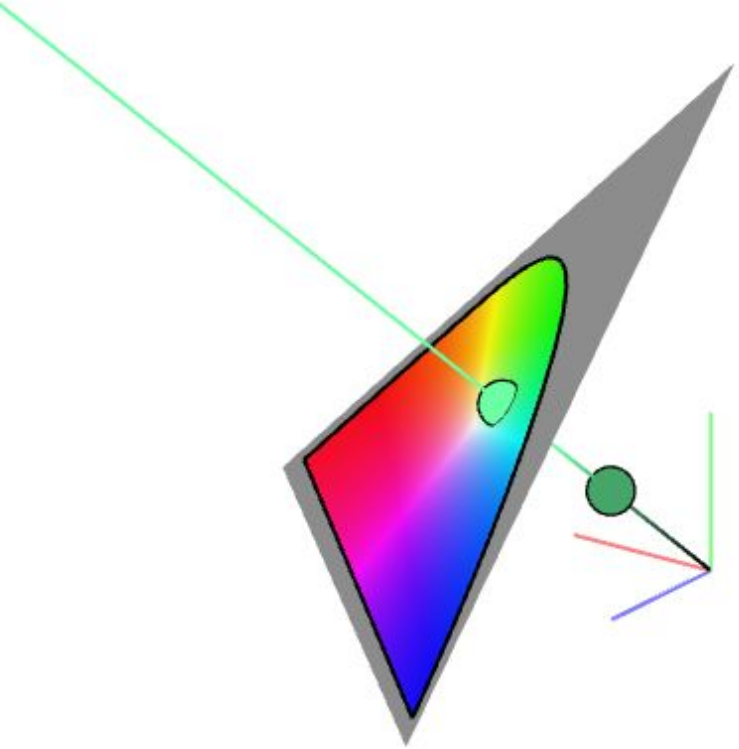


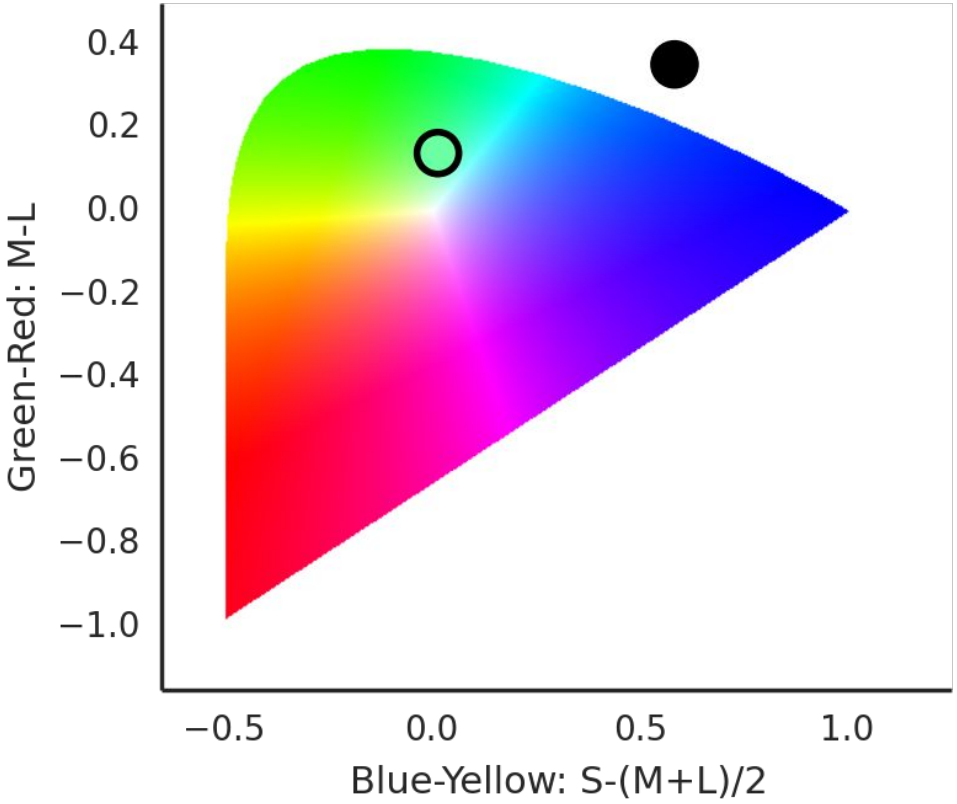







0.4
0.2
0.0
−0.2
−0.4
−0.6
−0.8
−1.0
Green-Red: M-L
−0.5
0.0
0.5
1.0
Blue-Yellow: S-(M+L)/2




0.4
0.2
0.0
−0.2
−0.4
−0.6
−0.8
−1.0
Green-Red: M-L
−0.5
0.0
0.5
1.0
Blue-Yellow: S-(M+L)/2

Visualizing High Dimensional Data

(in another context now gone horribly wrong)

Senate Data from 2008

Available online

Name	*Motion 1*	*Motion 2*	*Motion 3*	*Motion 4*	...
Alexander	Yea	Yea	Yea	Nay	
Biden	Yea	Nay	Nay	Yea	
Bond	Yea	Abstain	Nay	Nay	
Boxer	Yea	Yea	Yea	Abstain	
...					...

Senate Data from 2008

Available online

Name	*Motion 1*	*Motion 2*	*Motion 3*	*Motion 4*	...
Alexander	1	1	1	-1	
Biden	1	-1	-1	1	
Bond	1	0	-1	-1	
Boxer	1	1	1	0	
...					...

Notebook Example