Efficient Coding vs. Efficient Reconstruction in the Visual Cortex

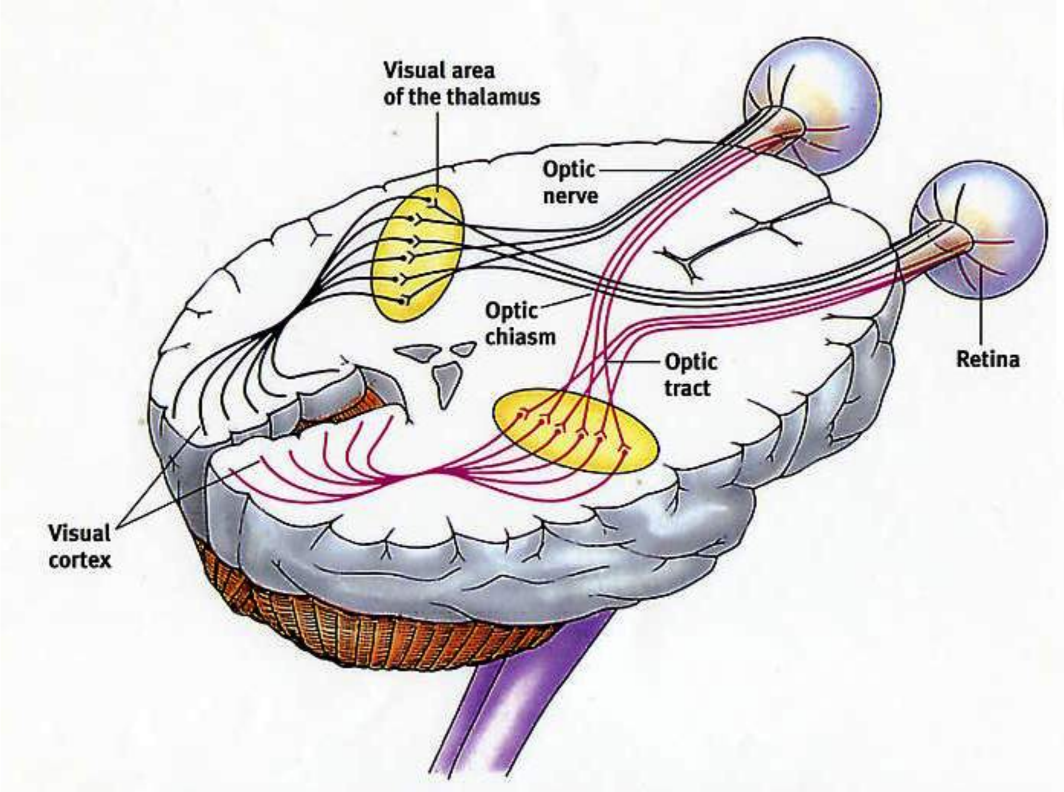
Yacov Hel-Or

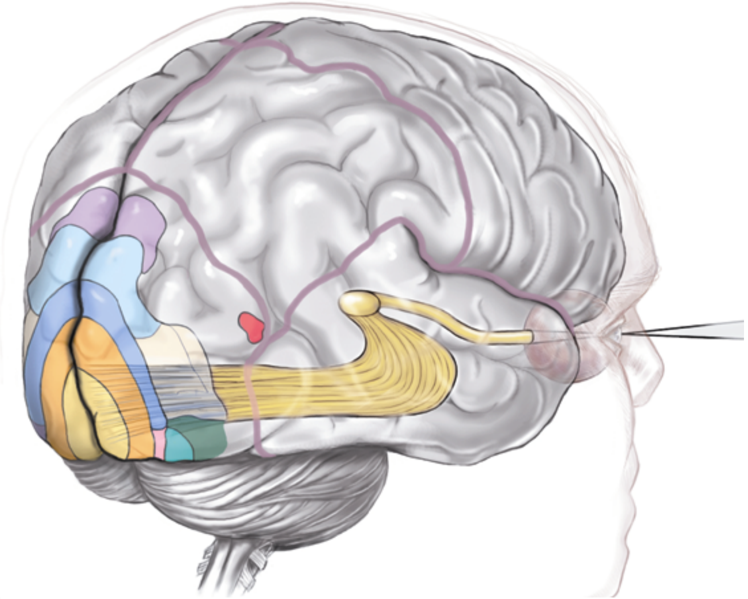
Efi Arazi School of Computer Science

The Interdisciplinary Center

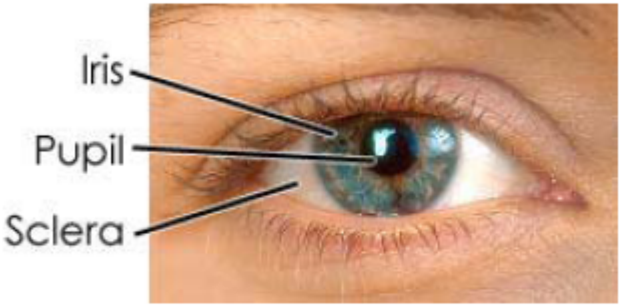
Herzliya, Israel

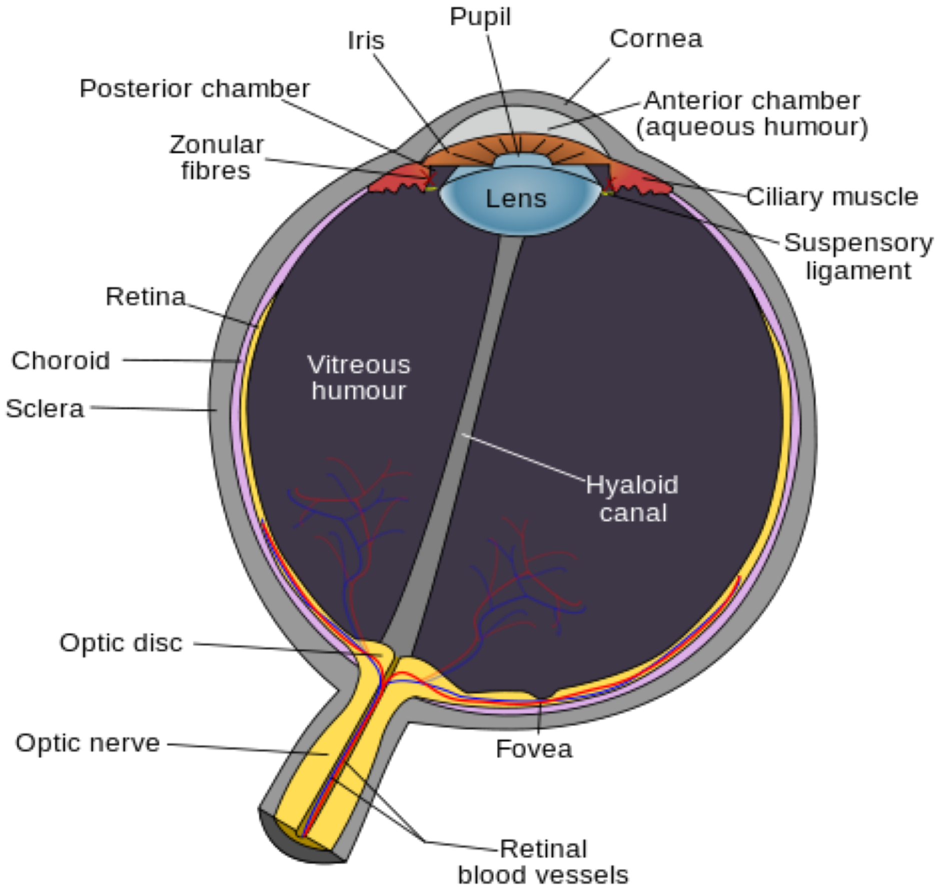
Some Trivial Facts about the Visual Pathway





The Human Eye





- Cornea - קרנית
- Pupil - אישון
- Iris - קשתית
- Retina - רשתית

The Retina

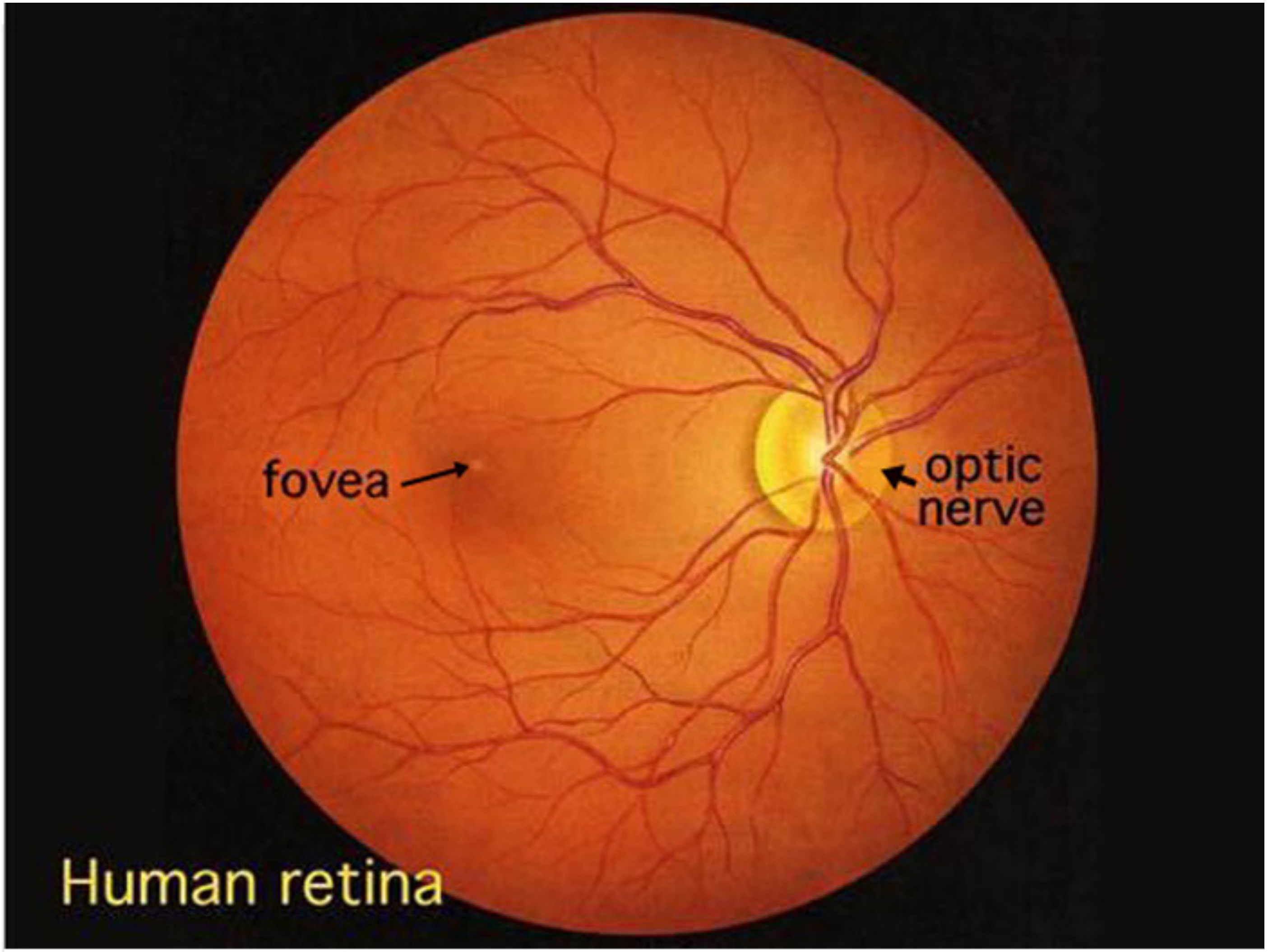
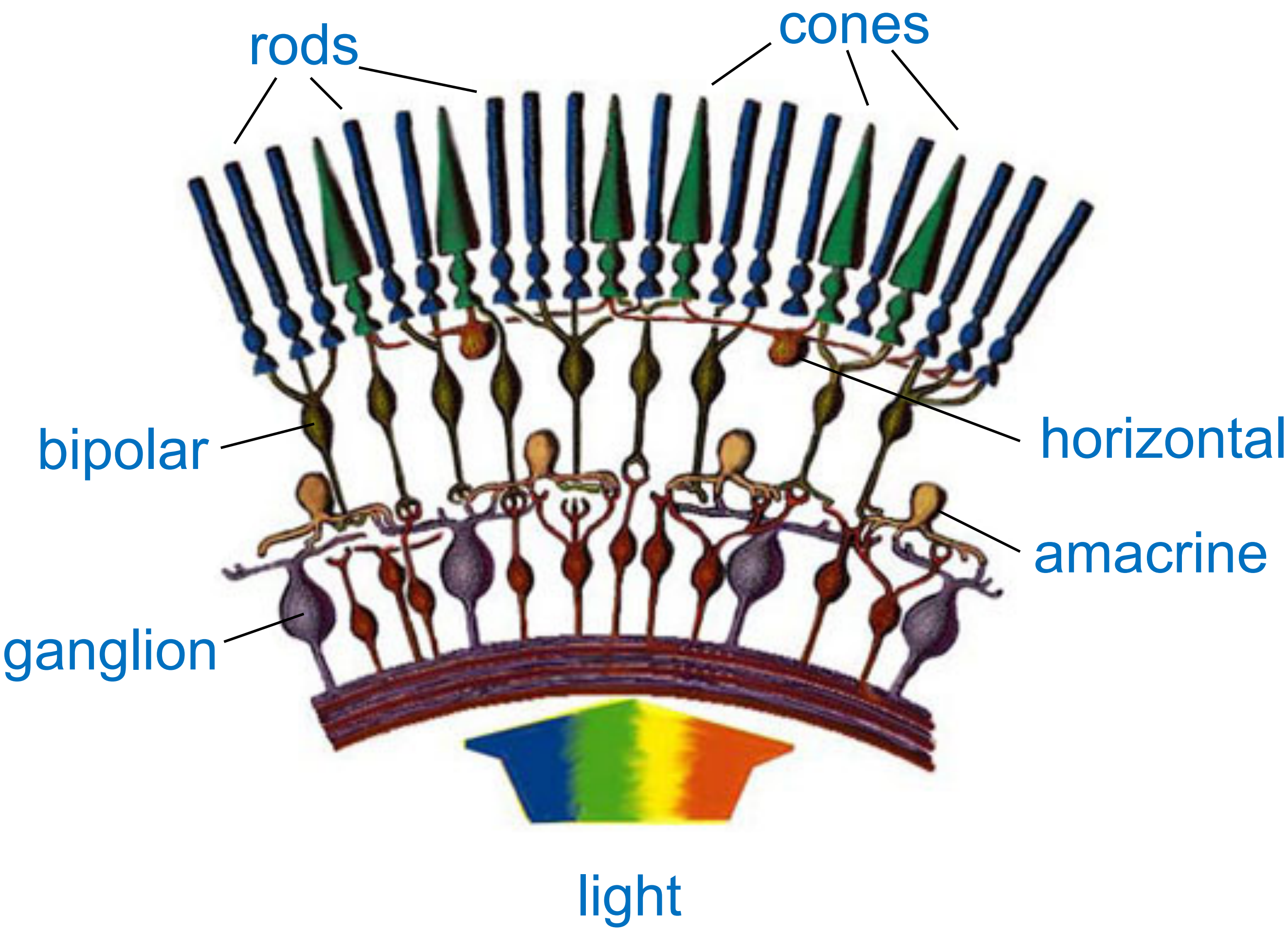


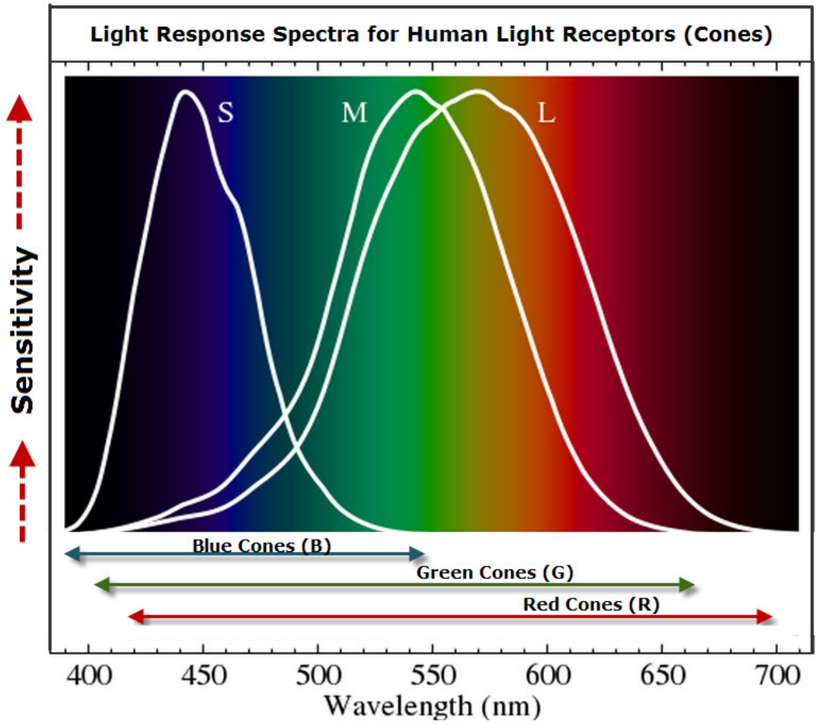
Fig. 1. Human retina as seen through an opthalmoscope.

- The retina contains two types of photo-receptors:
  - **Cones**: Photopic vision, can perceive color tones
  - **Rods**: Scotopic vision, can perceive brightness only



The Cones

- Three types of sensors: L, M, and S, each with different photo-pigment, composing the trichromatic color vision
- 6-7 million cone receptors, located primarily in the central portion of the retina

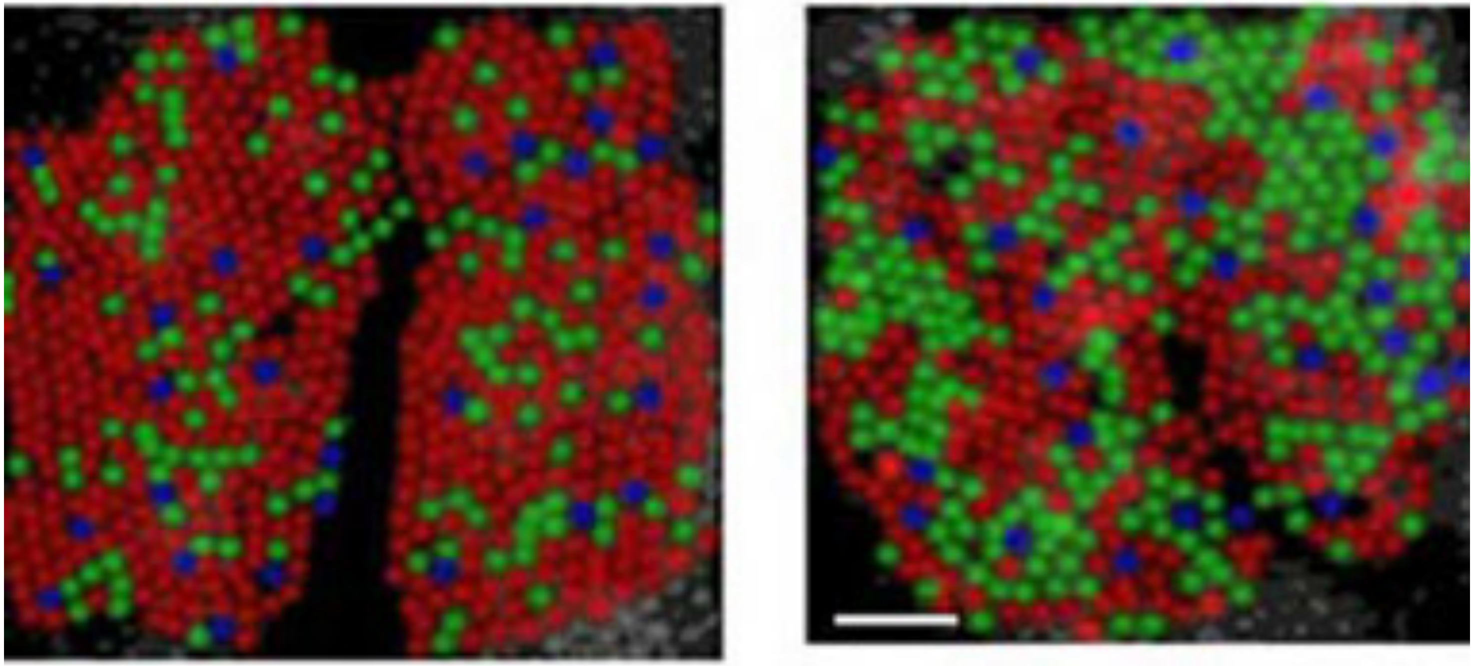


A side note:

- Humans and some monkeys have three types of cones (trichromatic vision); most other mammals have two types of cones (dichromatic vision).
- Marine mammals have one type of cone.
- Most birds and fish have four types.

Phenomena 1:

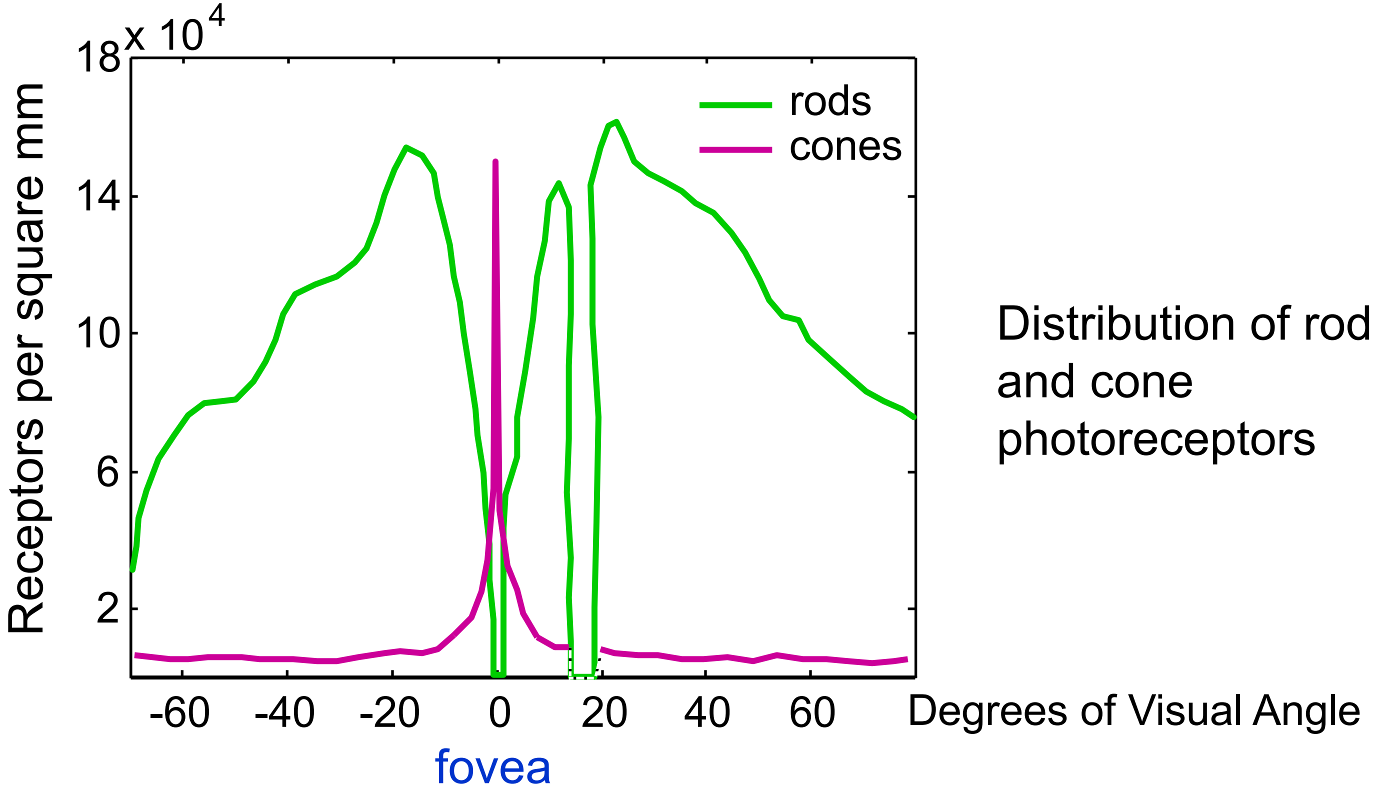
- Ratio of L to M to S cones is approx. 75:20:5
- Almost no S cones around the fovea



Cone Receptor Mosaic (from Roorda and Williams, 1999)

Cone Distribution:

- **L-cones** (Red) form about ~65% of the cones in the retina .
- **M-cones** (green) form about ~30% of the cones.
- **S-cones** (blue) form about ~2-5% of the cones



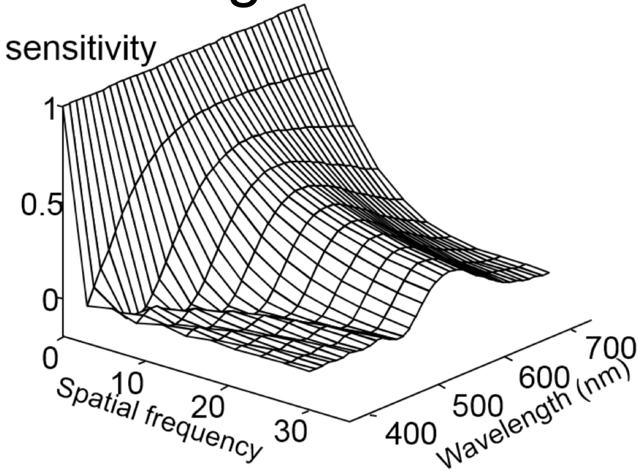
Distribution of rod and cone photoreceptors

Question 1:

Why is the distribution of L M S cones not uniform?

Common Answer 1:

Chromatic aberration of the lens, blurs the range of the S spectrum.



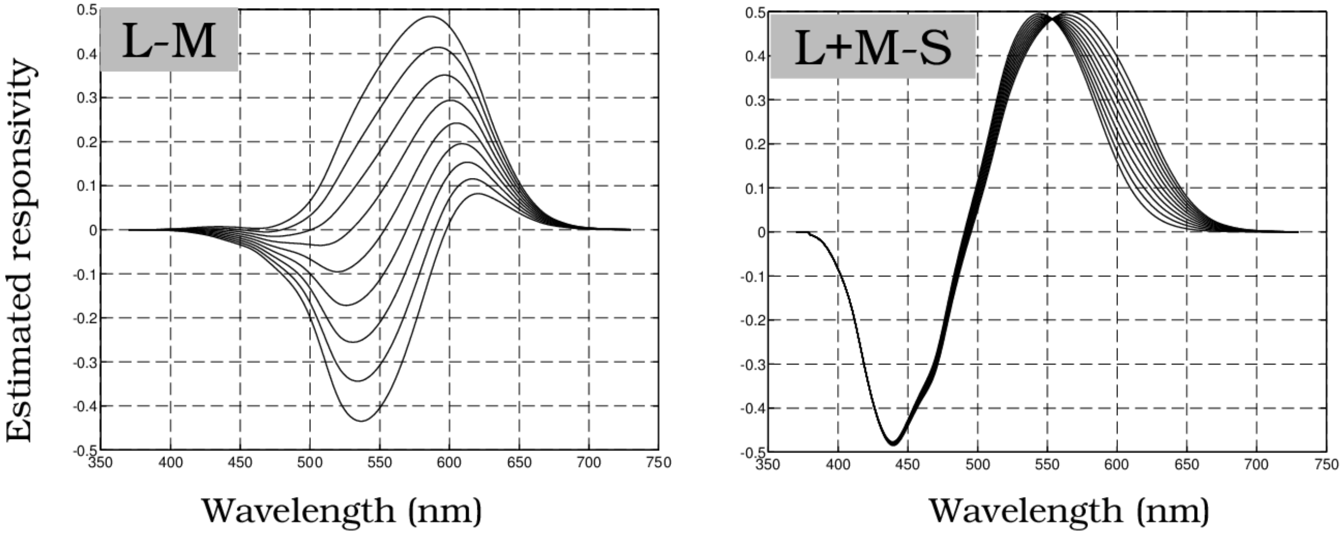
Common answer 2:

Blue colors are commonly smooth

Common answer 3: Evolution (no blue in *Odyssey* , *the Iliad,* and in the *Bible*).

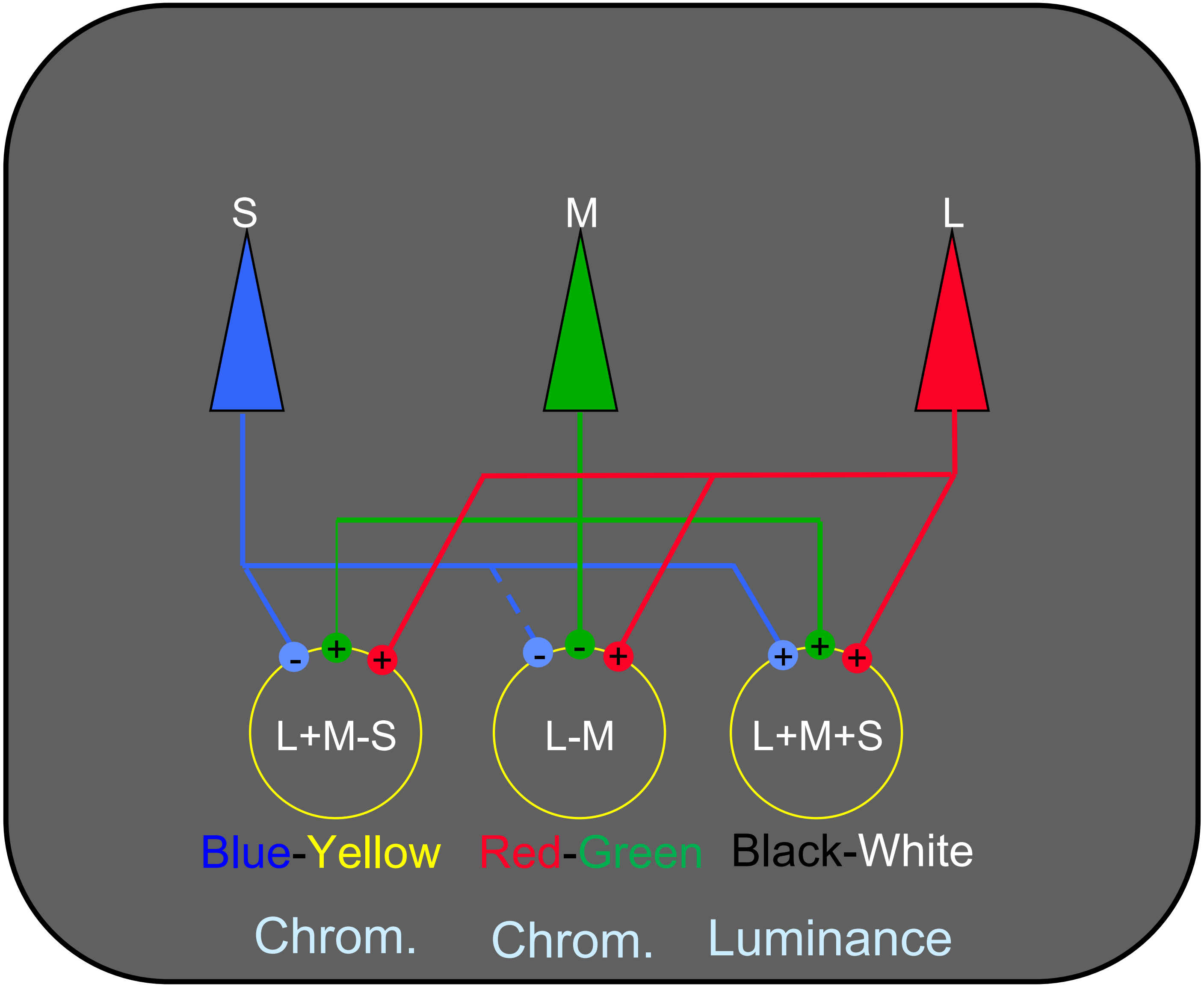
Phenomena 2:

Neurons in the visual cortex are sensitive to opponent signals (luminance/chrominance)



Derrington (1984)

Opponent Cells: possible neural connections



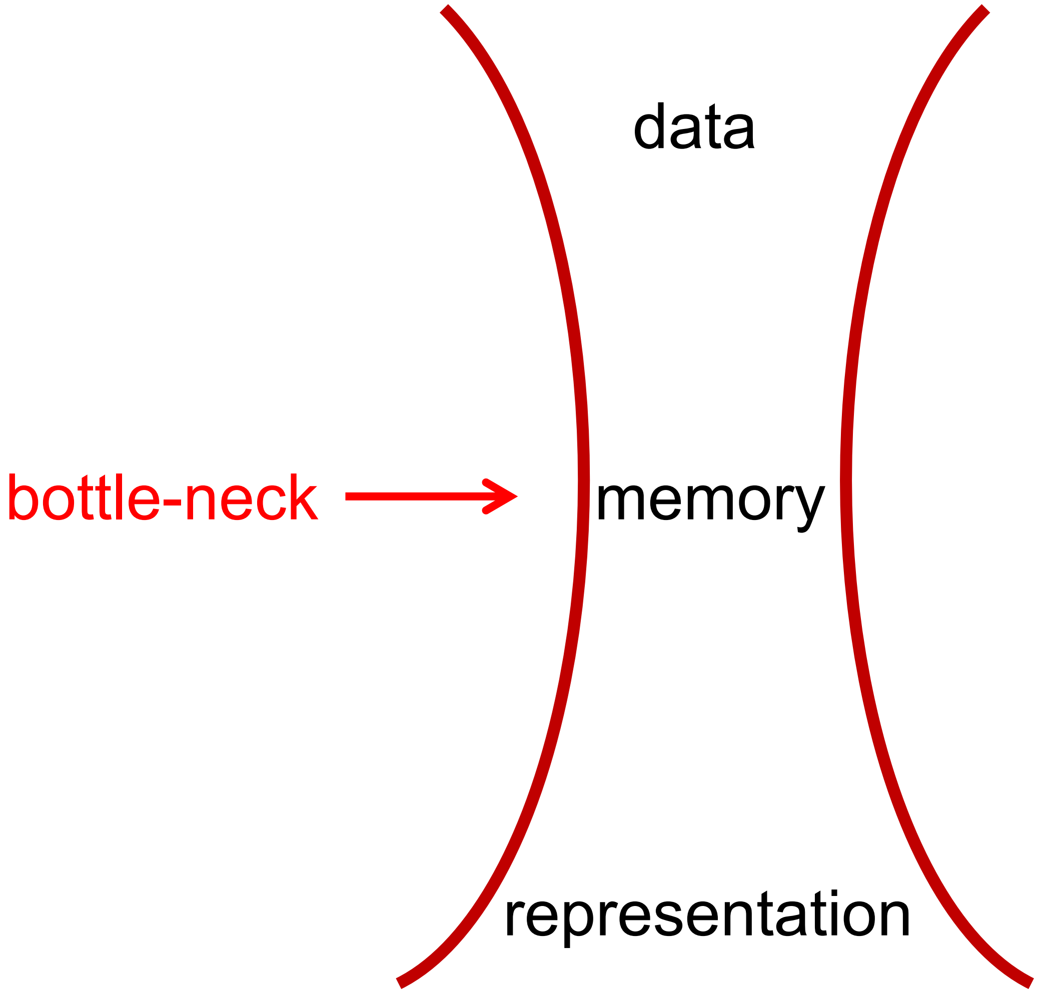
Question 2:

Why does the HVS encode color information in opponent space?

Common Answer:

“Efficient Coding“ – In order to reduce data redundancies opponent basis de-correlates color information (Barlow 89, Field 87) .

Efficient Coding



data

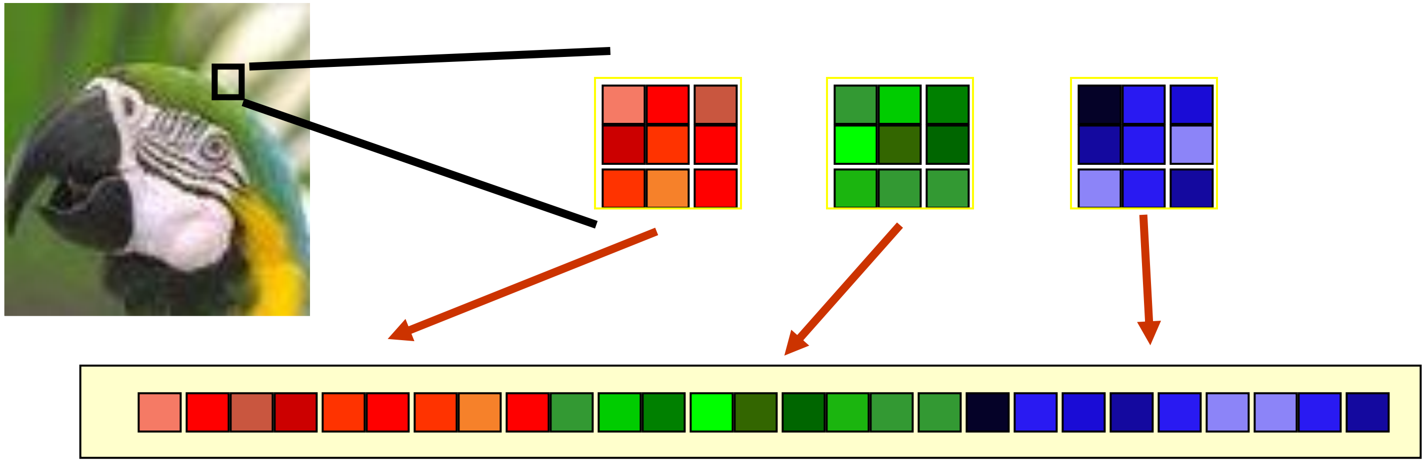
bottle-neck → memory

representation

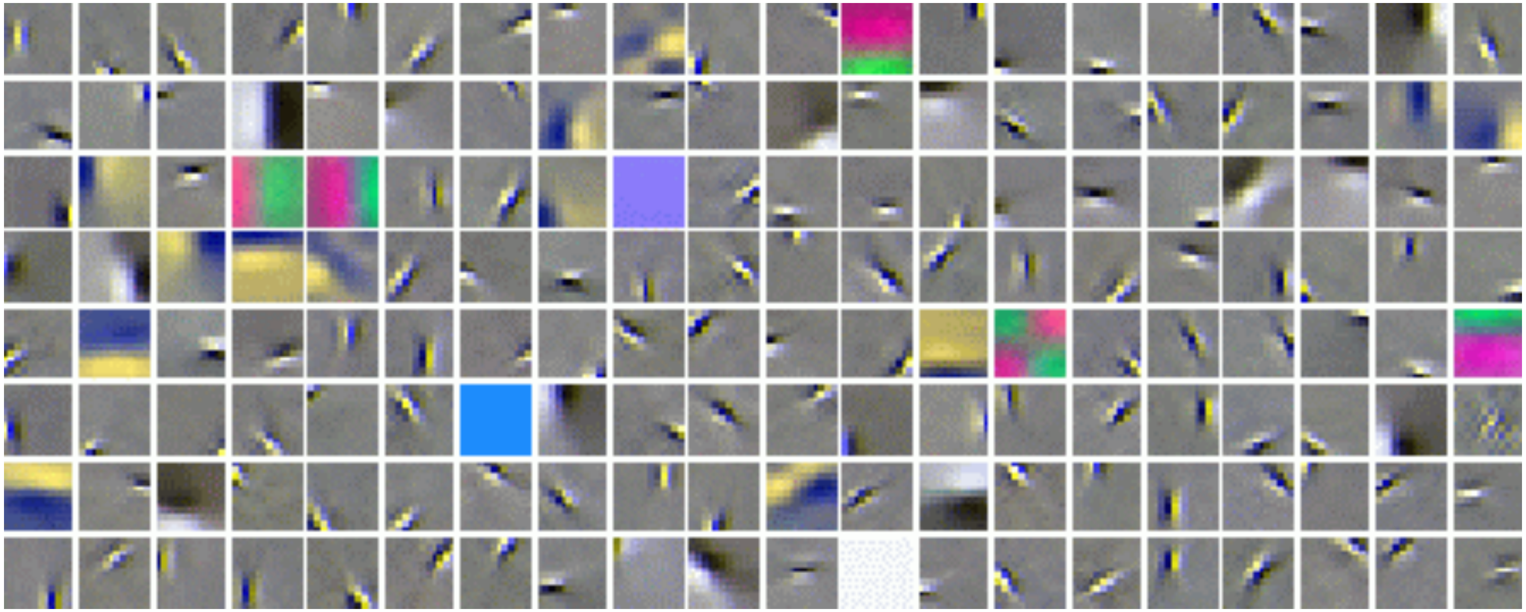
Efficient Coding

Method:

- Collect spatio-chromatic data from natural images.
- Whiten the data by projecting onto the principal components of PCA or ICA.



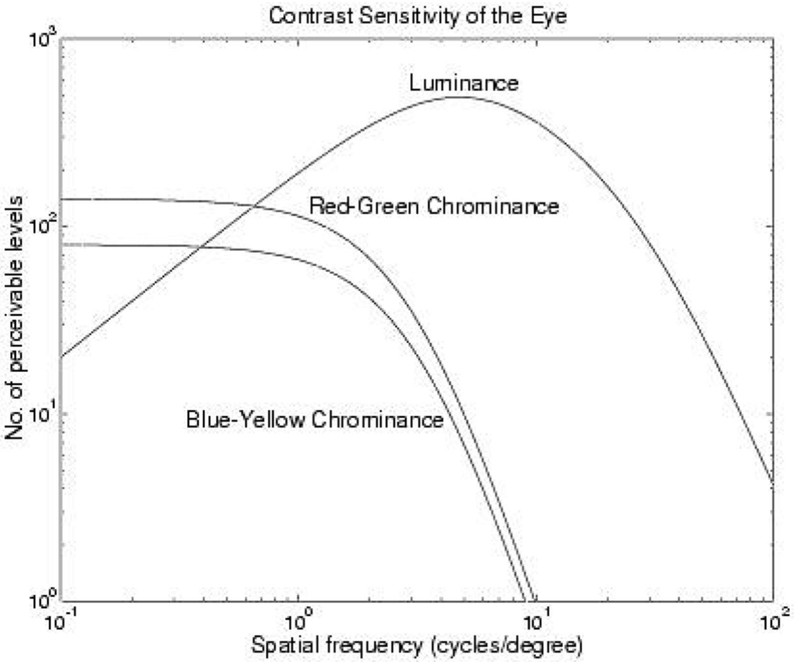
Efficient Coding using ICA Hoyer & Hyvärinen 2000



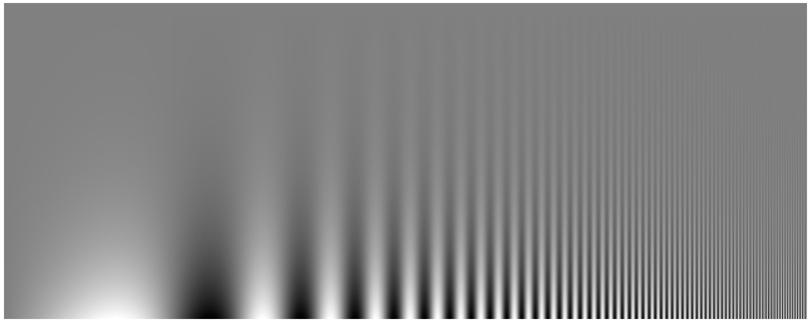
- Oriented, localized, and band passed basis
- Luminance/Chrominance arrangement (B-Y & R-G)
- High freq. for luminance basis and low freq. for chrominance
- Fewer number of chromatic basis

Phenomena 3:

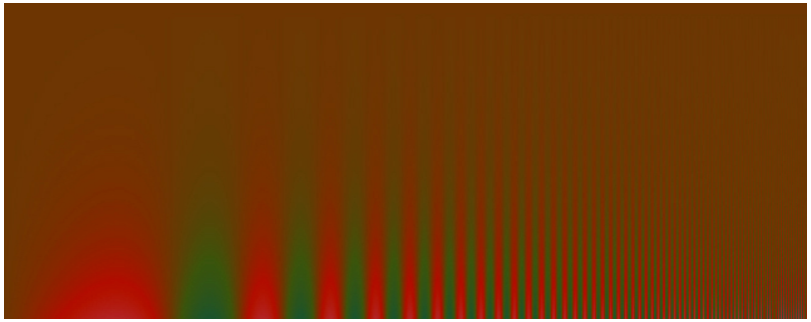
- The HVS is more sensitive to high-frequencies in the luminance channel than in the chrominance



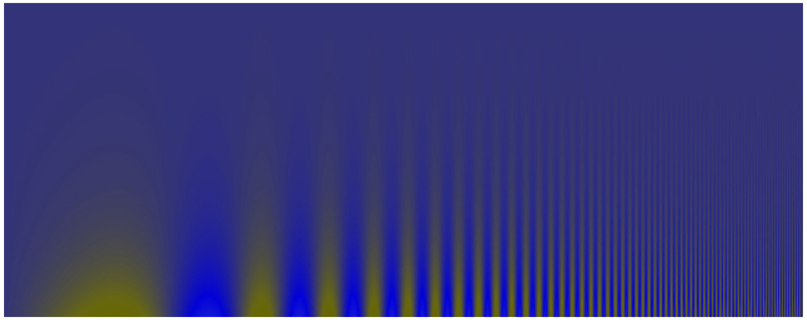
Luminance:



Red-Green:



Blue-Yellow:



Original Image

After blurring the two chrominance bands

After blurring the luminance band

Question 3:

Why is the HVS less sensitive to high spatial frequencies in the chrominance channels?

Common Answer: Efficient Coding → Chrominance cells are tuned for low spatial freq. and luminance cells for high spatial freq.

Efficient Coding and the HVS

Efficient coding agrees with the characteristics of the HVS

Data Whitening (PCA & ICA)	**H.V.S.**
Luminance/Chrominance arrangement of basis vectors (B-Y & R-G)	Luminance/chrominance pathways in the visual cortex
Spatial basis vectors are oriented, localized, and band passed.	Resembles the Simple/Complex cells RFs
High spatial freq. for luminance basis vectors and low freq. for chrominance basis.	HVS is more spatially sensitive to luminance data.
Fewer number of chromatic basis vectors.	In accord with RFs in the HVS.

Buchsbaum & Gottschalk 83, Attick & Redlich 92, Olshausen & Field 96, Ruderman et.al. 98, Hoyer and Hyvärinen 2000, and more...

Summary:

1. The retina contains many more L, M cones than S cones.
2. The visual pathway encodes color information using luminance/chrominance channels.
3. The HVS is insensitive to high spatial-frequencies in the chrominance channel.

Our claim:

- All the above are related and stems from the statistical properties of color images and the ***shortage of sensory interface***.

The statistical properties of color images

The statistical properties of color images

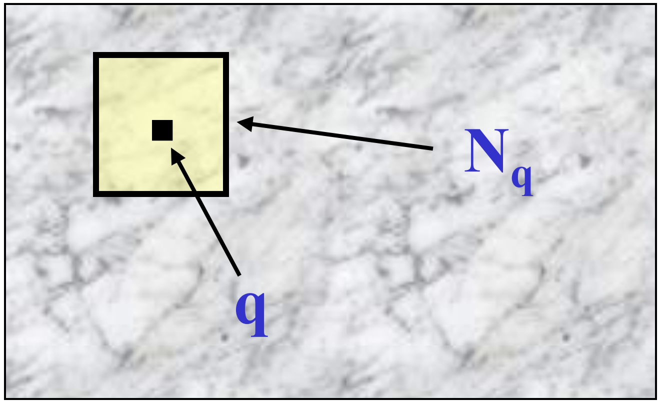
- Given a color image f modeling the entire joint probability P_F is impractical.
- In order to build a *useful* model we must reduce the dimensionality of the problem.

- Common approaches to image modeling use 2 types of reductions:
  - Reduction in the Spatial domain.
  - Projection onto *informative* subspaces.

Reduction in the Spatial Domain

- **A reasonable assumption**: A natural image can be viewed as a realization of a Markov Random Field:
  - A large enough neighborhood of an image pixel completely characterizes its p.d.f:

$$P\left(q|\,N_q\right)=P\left(q\mid p,p\neq q\right)$$

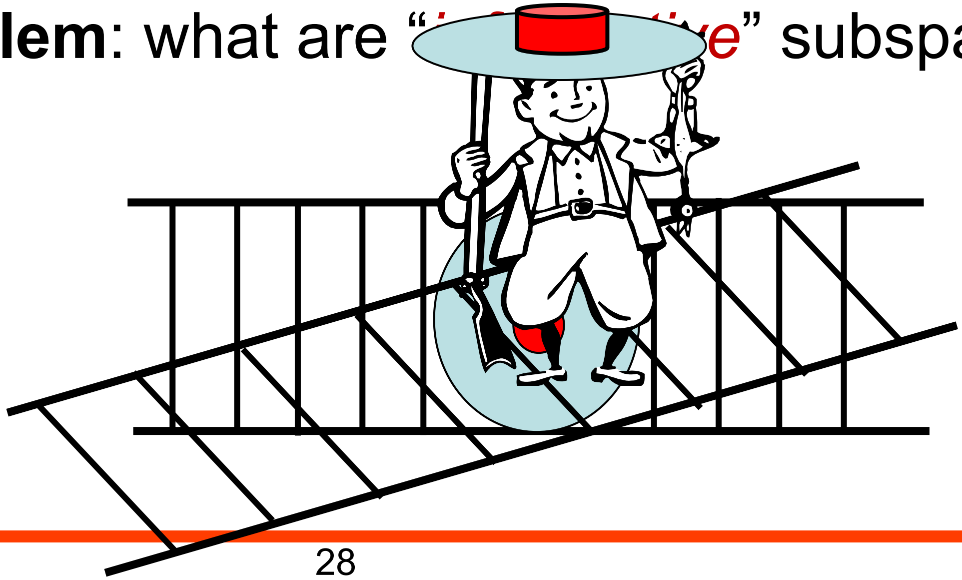


  - This p.d.f is similar for all pixels (the homogeneity property of images)

$$P\left(q,N_q\right)$$

Projection into Informative Subspace

- Modeling the Statistical properties of $k \times k \times 3$ spatio-chromatic patch is complicated, but an approximation can be inferred using marginal statistics in some projected subspace.
- Subspaces should be chosen such that "*informative*" information will not be lost.
- **A crucial problem**: what are "*informative*" subspaces?



Informative Subspaces in Color Images

- Suggested approach: Choose subspaces in which the correlations between the color bands will be maximized.

The Canonical Correlation Analysis (CCA)

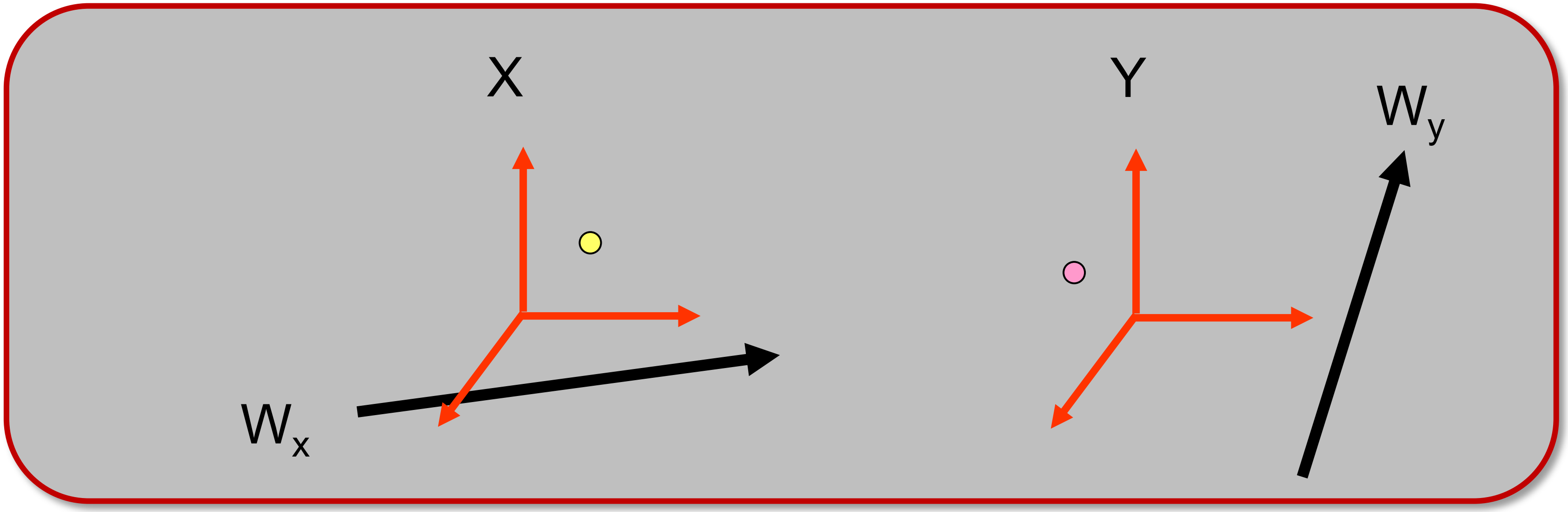
finds such subspaces.

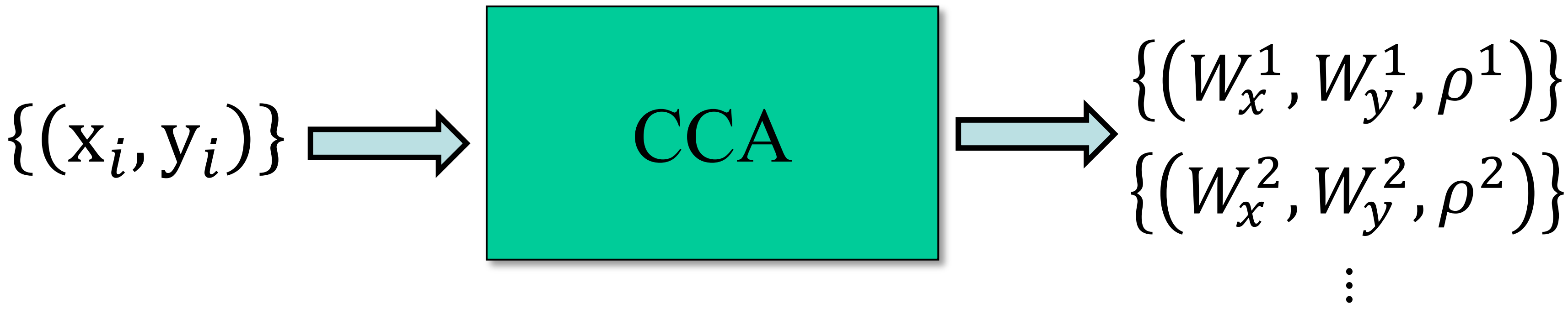
The Canonical Correlation Analysis (CCA)

- Assume two multidimensional random variables: x and y .
- We are looking for two projection vectors W_x and W_y such that the correlation between $x' = x^T W_x$ and $y' = y^T W_y$ is maximized:

$$\rho(w_x, w_y) = \frac{\mathrm{cov}(x', y')}{\sigma_{x'}\sigma_{y'}}$$

Taken from Kidron et. Al.





$\{(x_i, y_i)\}$

CCA

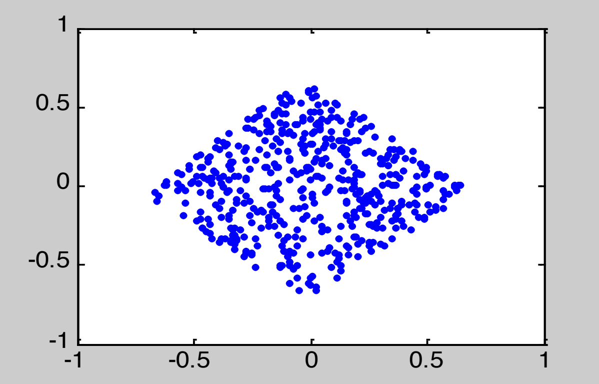
$\{(W_x^1, W_y^1, \rho^1)\}$

$\{(W_x^2, W_y^2, \rho^2)\}$

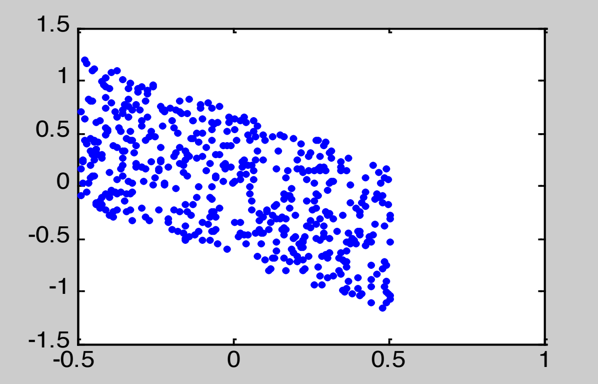
$\vdots$

- W_x^1 and W_y^1 define the directions with the maximal correlation ρ^1
- W_x^2 and W_y^2 are the second best directions, and so forth
- The set $\{(W_x^i, W_y^i)\}$ are the ***C.C. basis vectors***
- The corresponding $\{\rho^i\}$ are the ***canonical correlations***

Illustrating Example



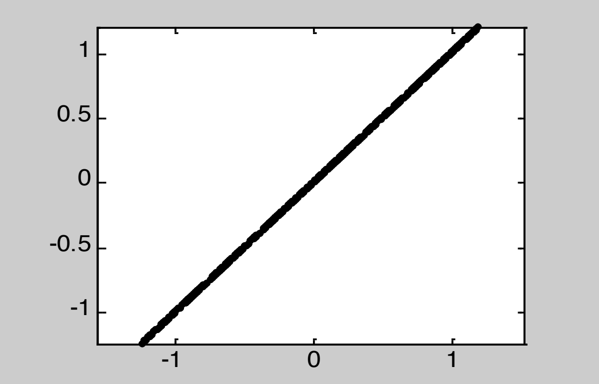
$x_1 - x_2 = 2y_1 + y_2$



$\mathbf{x}=(x_1,x_2)$

$\mathbf{y}=(y_1,y_2)$

x’



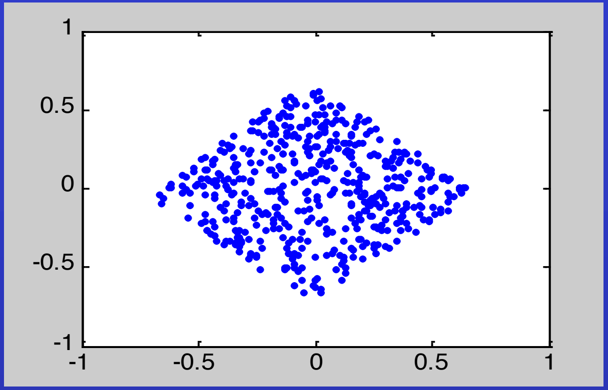
y’

$x' = (1, -1)\, \mathbf{x}$

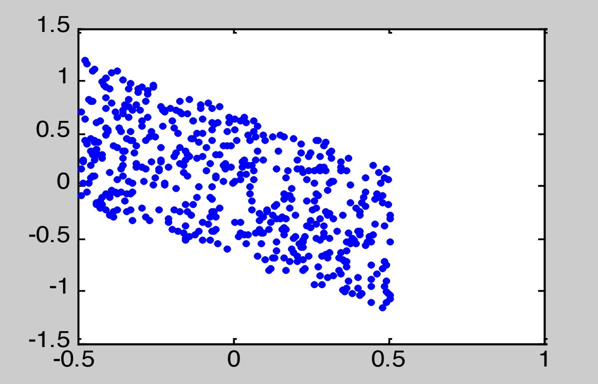
$y' = (2, 1)\, \mathbf{y}$

Canonical Correlations

Previous Example using PCA

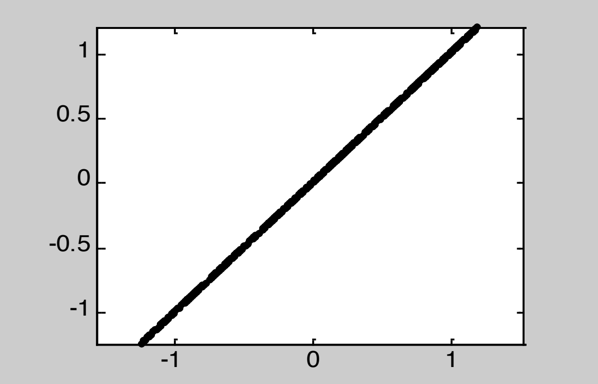


$\mathbf{x}=(x_1,x_2)$

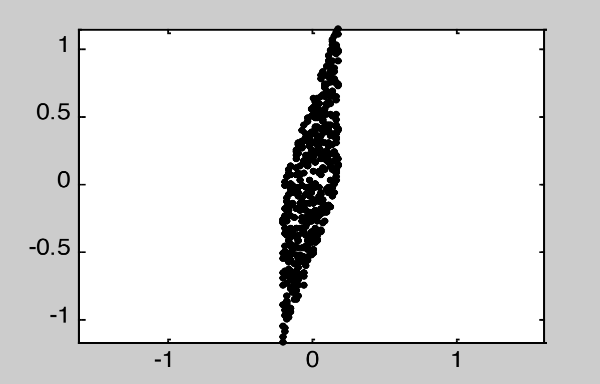


$\mathbf{y}=(y_1,y_2)$

$$x_1 - x_2 = 2y_1 + y_2$$



Canonical Correlations

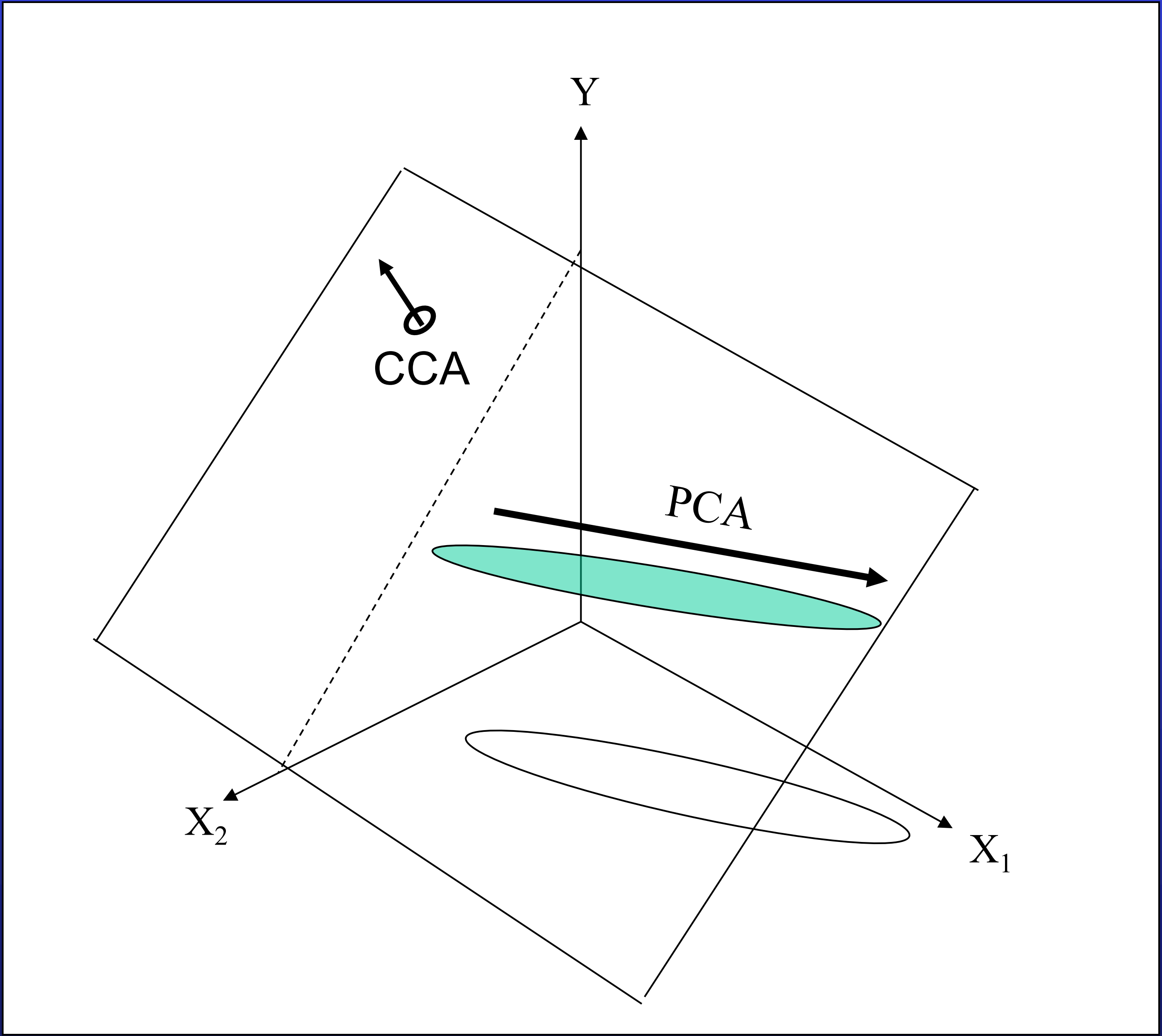


Principal Components

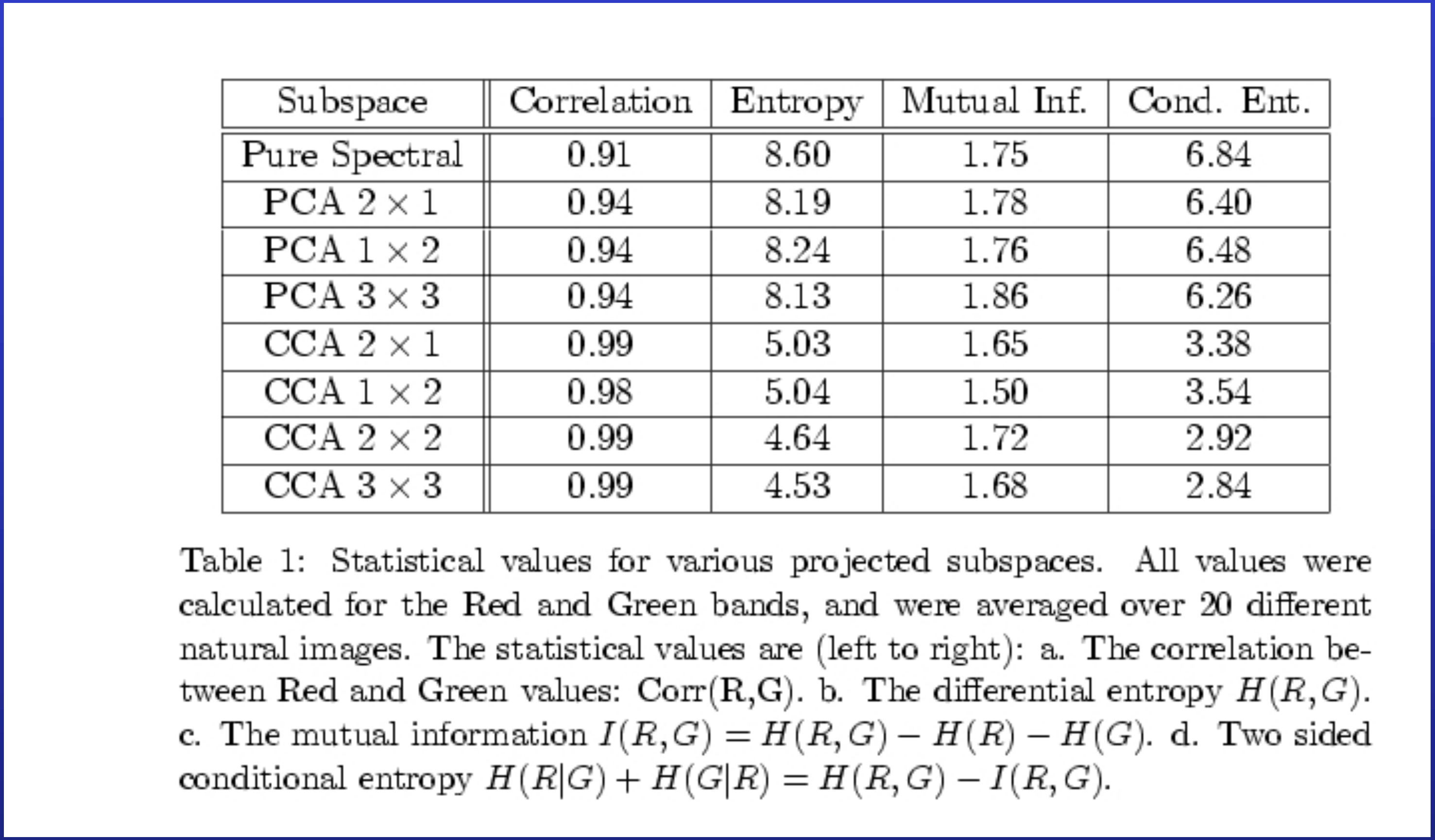
CCA V.S. PCA

$y=1-x_2+n$

Due to the *within* correlations of *x*, PCA fails to expose the mutual dependencies between x and y.



CCA V.S. PCA



Subspace	Correlation	Entropy	Mutual Inf.	Cond. Ent.
Pure Spectral	0.91	8.60	1.75	6.84
PCA 2 × 1	0.94	8.19	1.78	6.40
PCA 1 × 2	0.94	8.24	1.76	6.48
PCA 3 × 3	0.94	8.13	1.86	6.26
CCA 2 × 1	0.99	5.03	1.65	3.38
CCA 1 × 2	0.98	5.04	1.50	3.54
CCA 2 × 2	0.99	4.64	1.72	2.92
CCA 3 × 3	0.99	4.53	1.68	2.84

Table 1: Statistical values for various projected subspaces. All values were calculated for the Red and Green bands, and were averaged over 20 different natural images. The statistical values are (left to right): a. The correlation between Red and Green values: Corr(R,G). b. The differential entropy $H(R,G)$. c. The mutual information $I(R,G) = H(R,G) - H(R) - H(G)$. d. Two sided conditional entropy $H(R|G) + H(G|R) = H(R,G) - I(R,G)$.

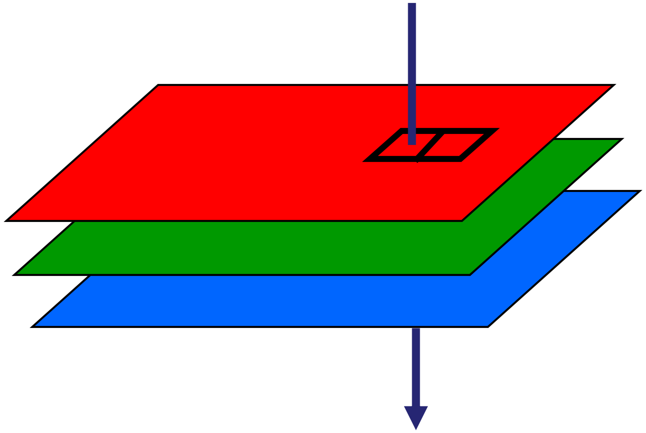
The CC of Color Images

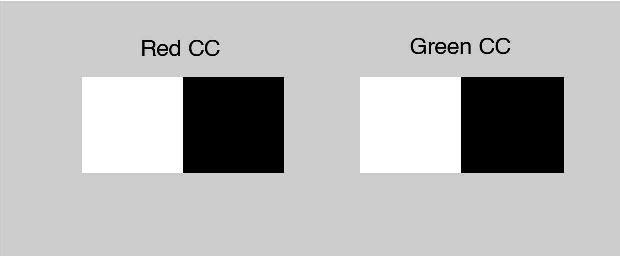
- In the following demonstrations we consider this image
- All values are presented in log(RGB) space

$$f(x,y)=\log\begin{pmatrix}R(x,y)\\G(x,y)\\B(x,y)\end{pmatrix}$$

The CC of 1x2 neighborhoods

- Applying the CCA over (R,G,B), where each variable is a 1x2 neighborhood, gives the following results:

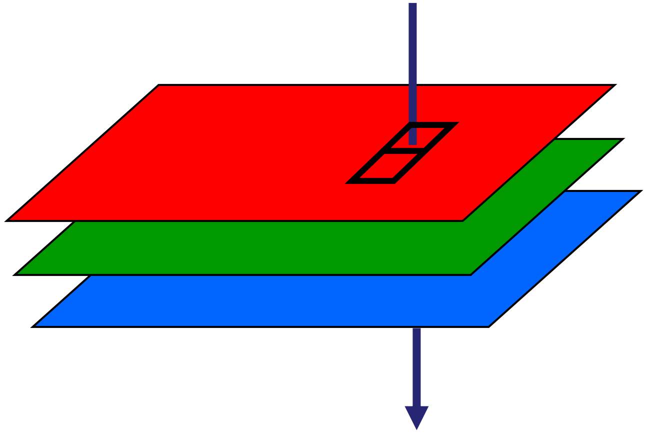


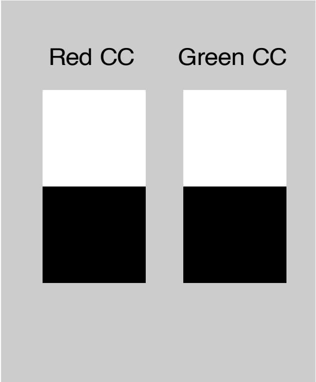


- The CC basis is composed of x-derivatives
- Similar results for each color pair

The CC of 2x1 neighborhoods

- Applying the CCA over (R,G,B), where each variable is a 2x1 neighborhood, gives the following results:

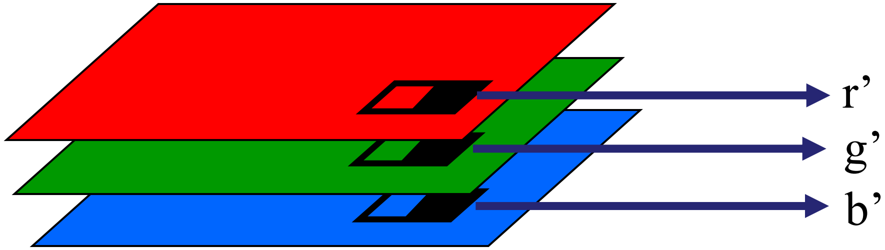




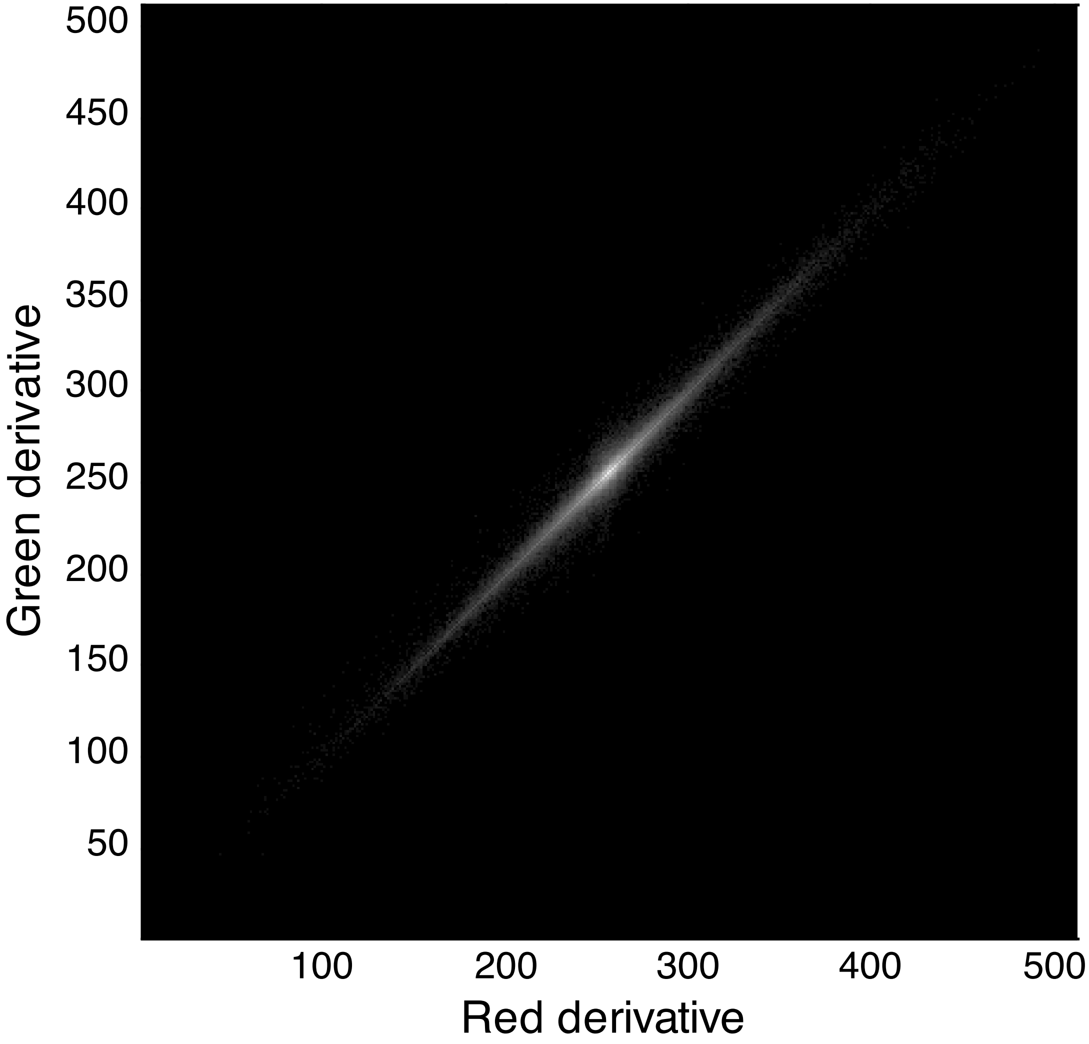
- The CC basis is composed of y-derivatives
- Similar results for each color pair

- The following joint histograms show the marginal p.d.f along the first CCA direction, for 1x2 neighborhoods:

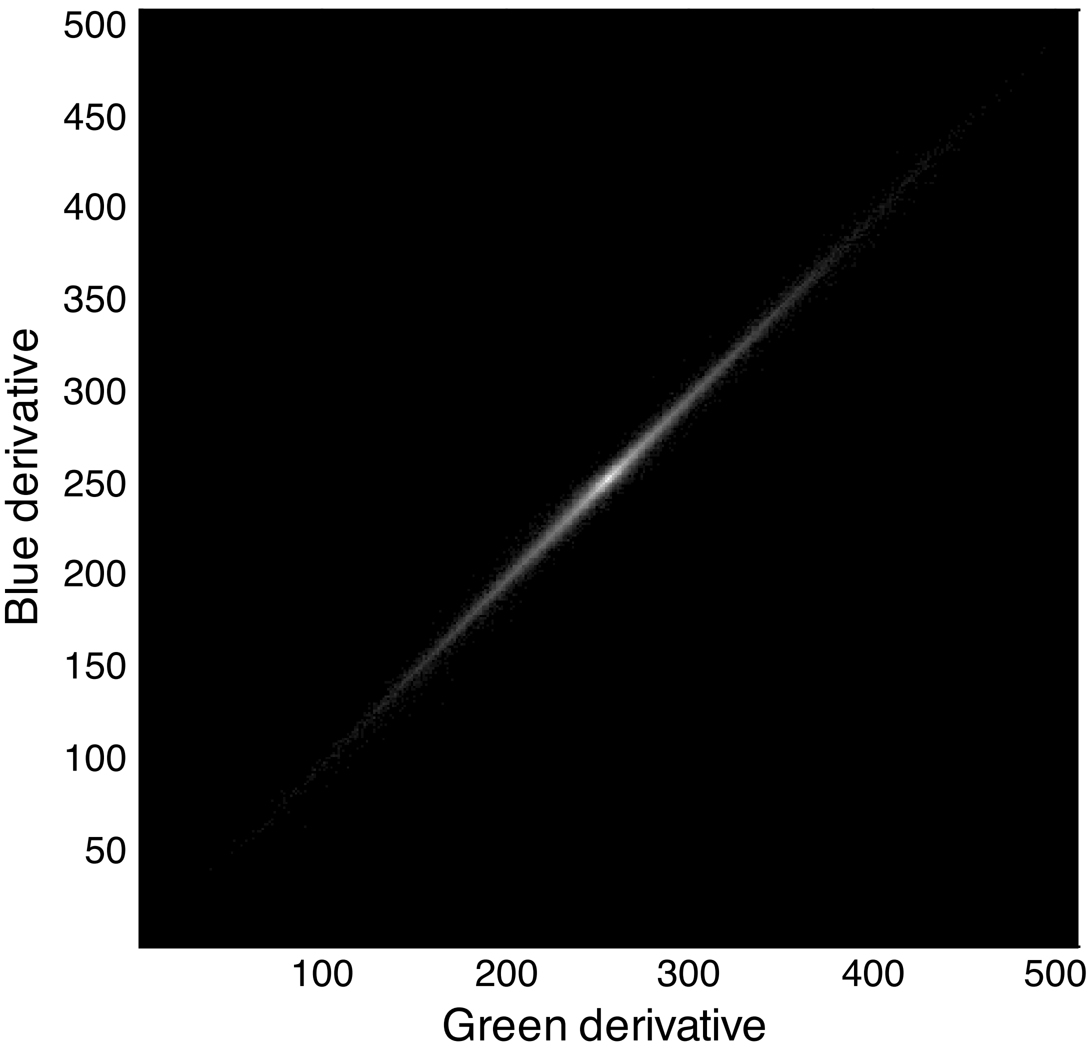
$$P(r',g',b') \propto H\left(\mathbf{r}(x,y)\cdot\mathbf{w}_R,\ \mathbf{g}(x,y)\cdot\mathbf{w}_G,\ \mathbf{b}(x,y)\cdot\mathbf{w}_B\right)$$

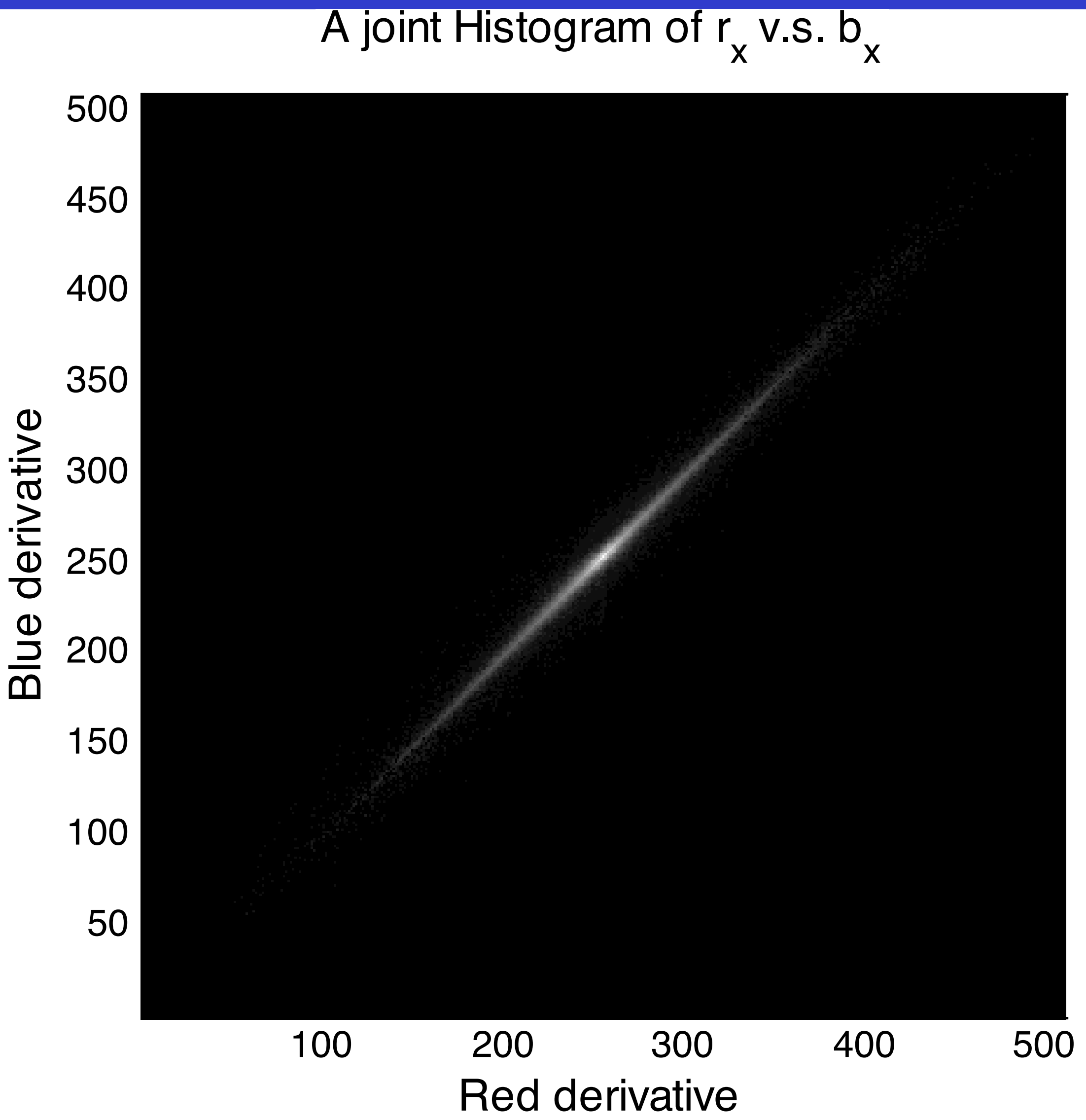


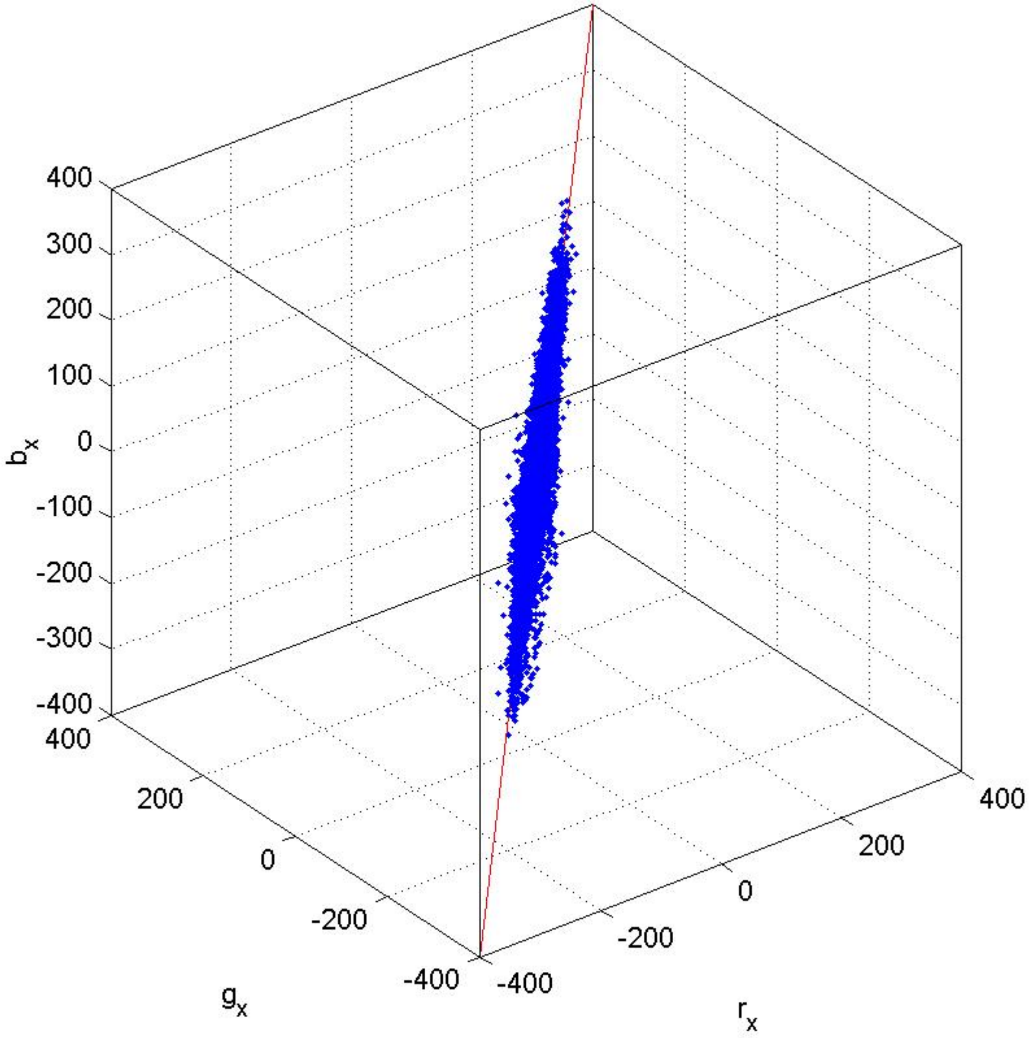
A joint Histogram of r_x v.s. g_x



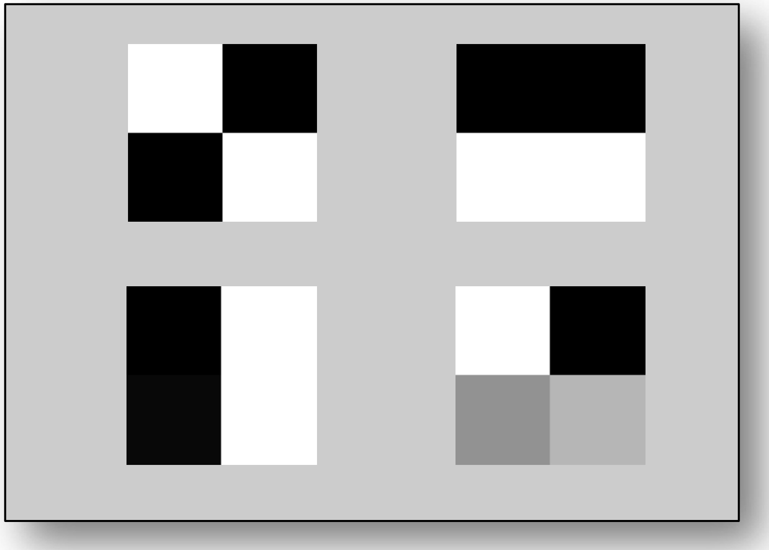
A joint Histogram of g_x v.s. b_x



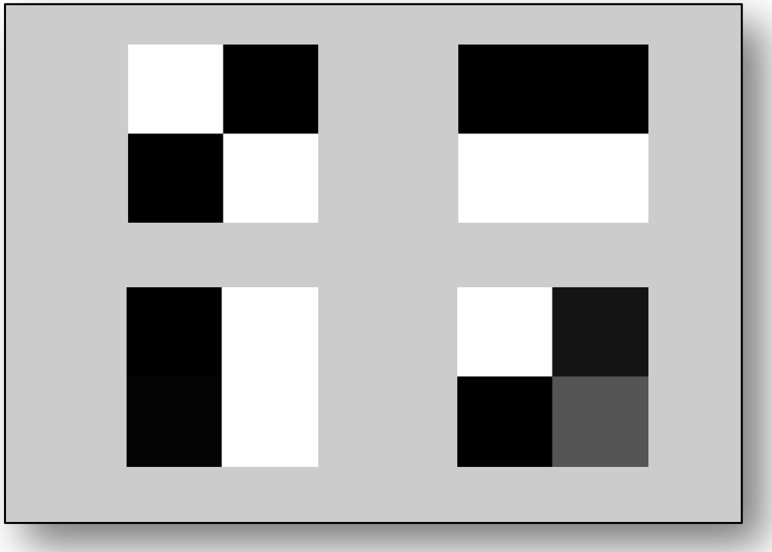

A joint Histogram of r_x v.s. b_x
Blue derivative
500
450
400
350
300
250
200
150
100
50
100
200
300
400
500
Red derivative


400
300
200
100
0
-100
-200
-300
-400
b_x
400
200
0
-200
-400
g_x
-400
-200
0
200
400
r_x

- Applying the CCA over (R,G,B) where each variable is a 2x2 neighborhood gives the following results:

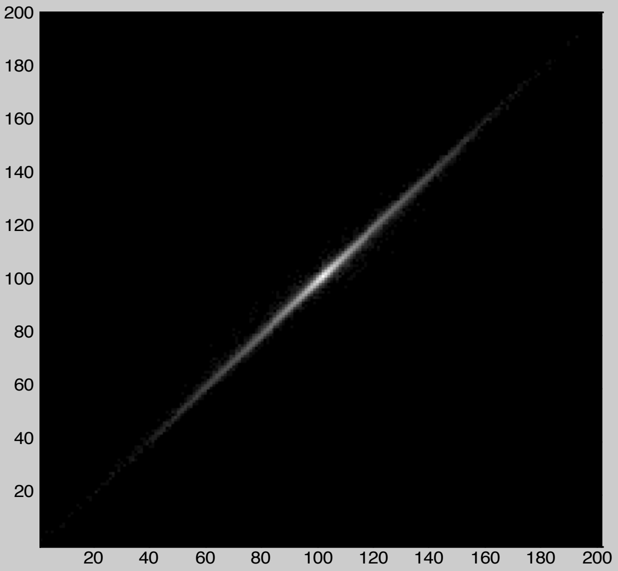


The 4 CC vectors of
the Red-Green plane

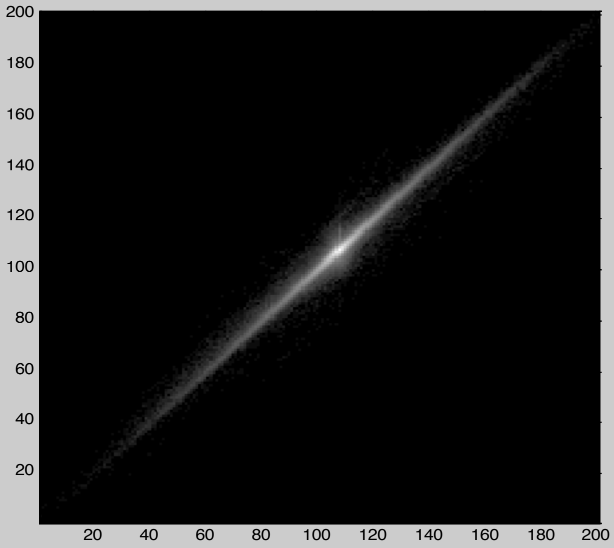


The 4 CC vectors of
the Green-Blue plane

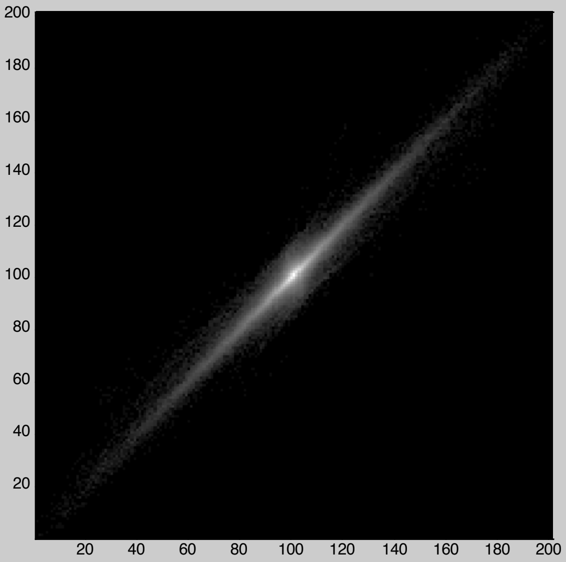
1st CC



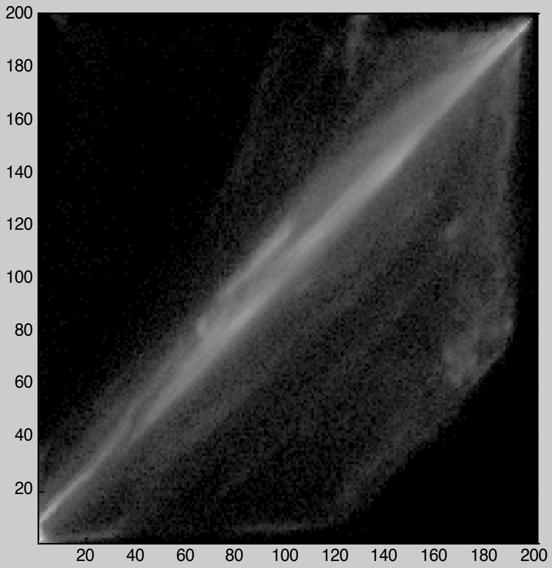
2nd CC

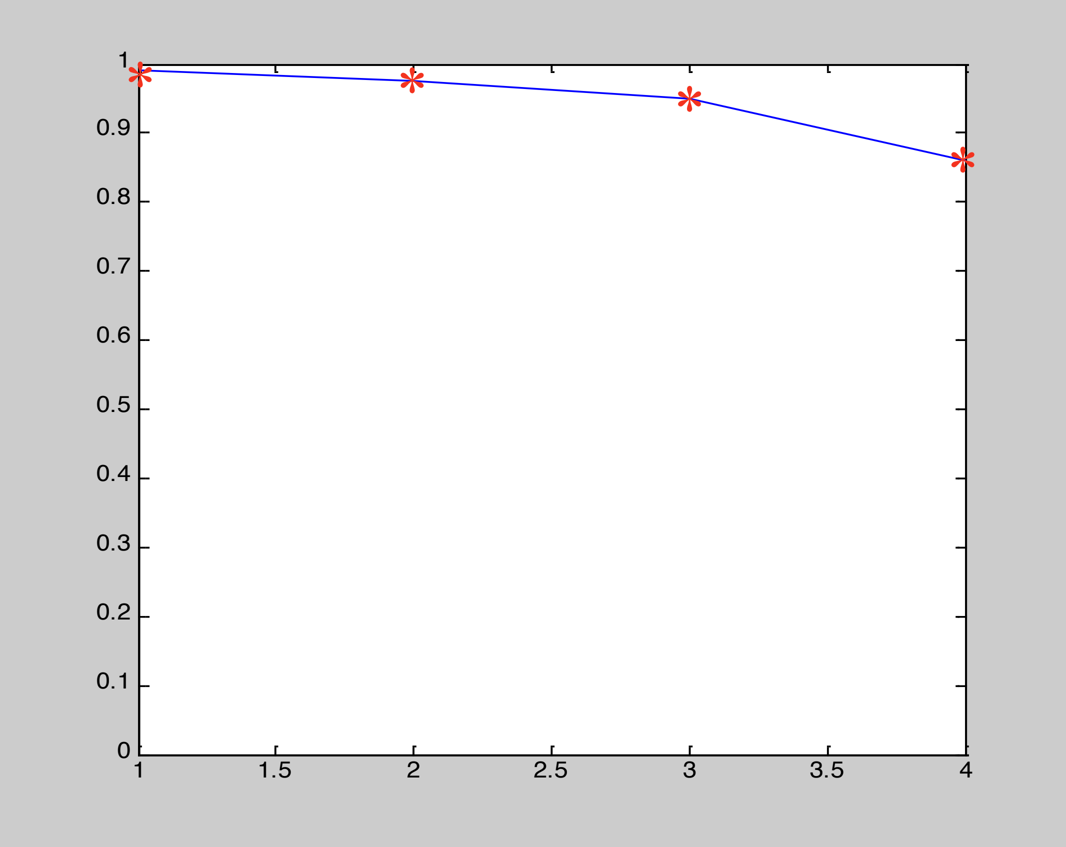


3rd CC



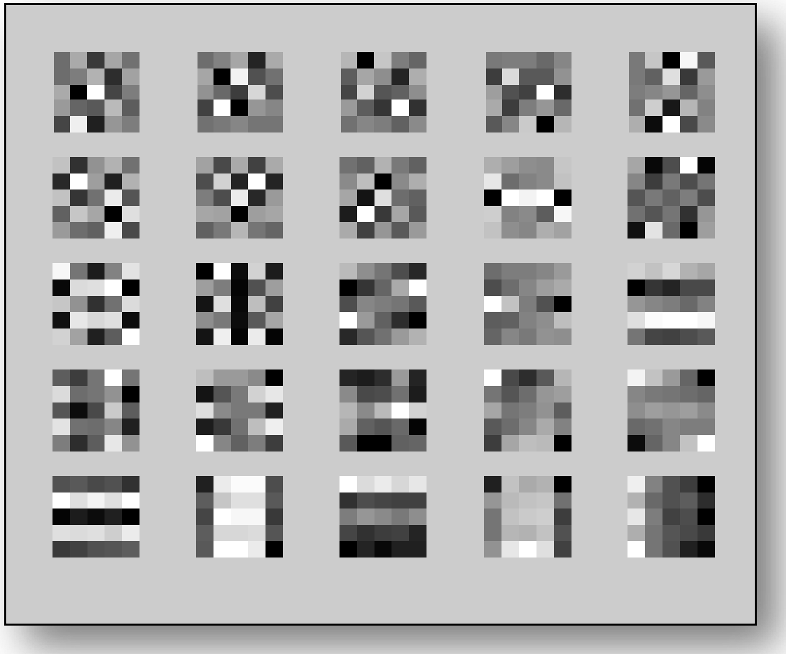
4th CC



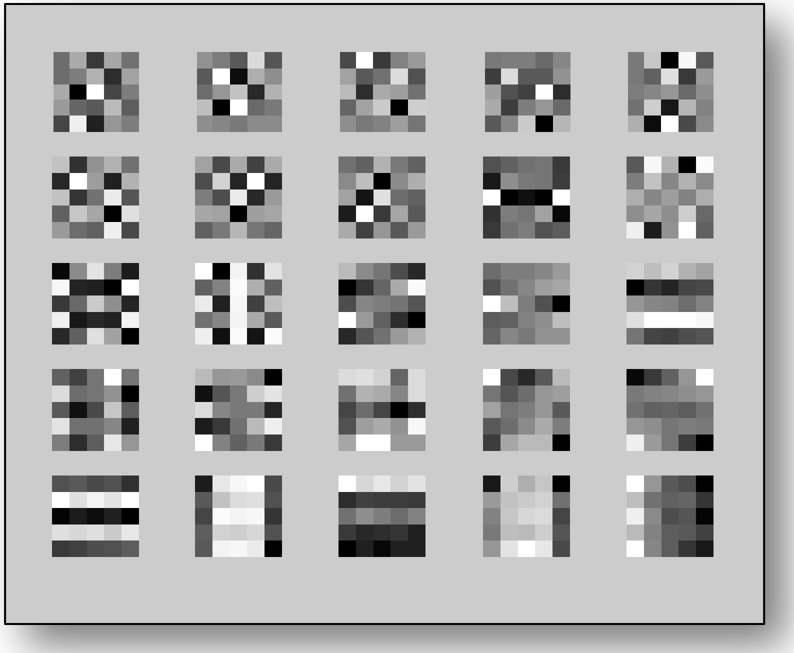


The 4 CC values for the CC vectors

- Applying the CCA over (R,G,B) where each variable is a 5x5 neighborhood gives the following results:

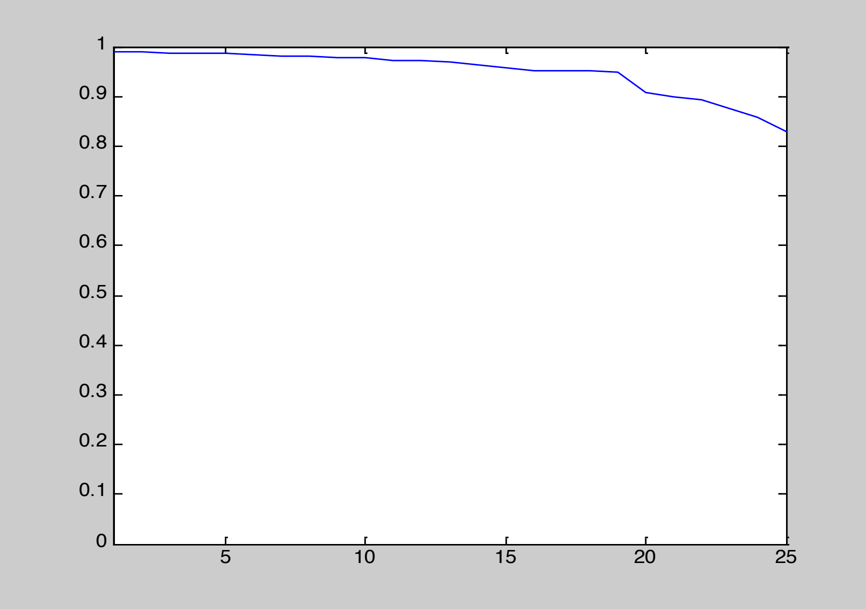


Red Plane

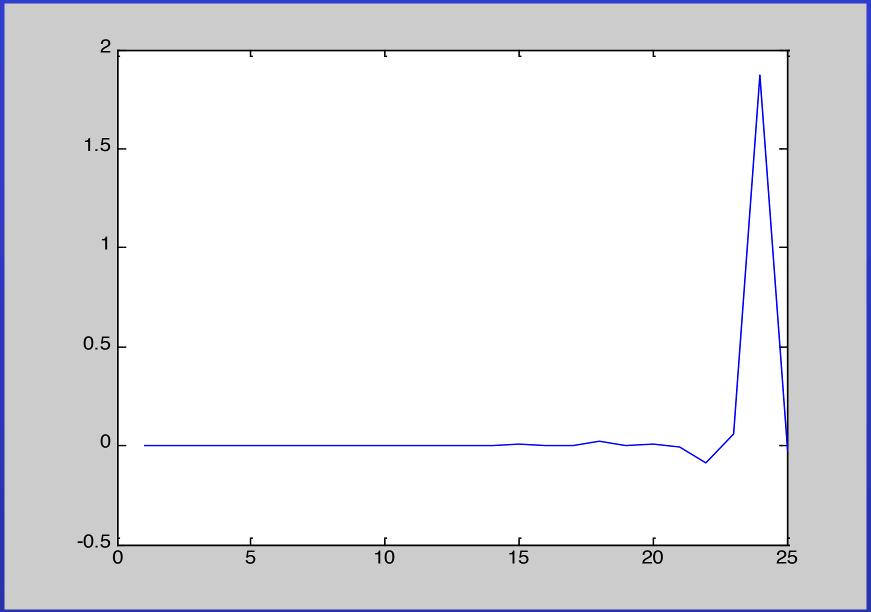


Green Plane

The 25 CC vectors



The 25 CC values for the CC vectors



The 25 DC values for the CC vectors

Observations

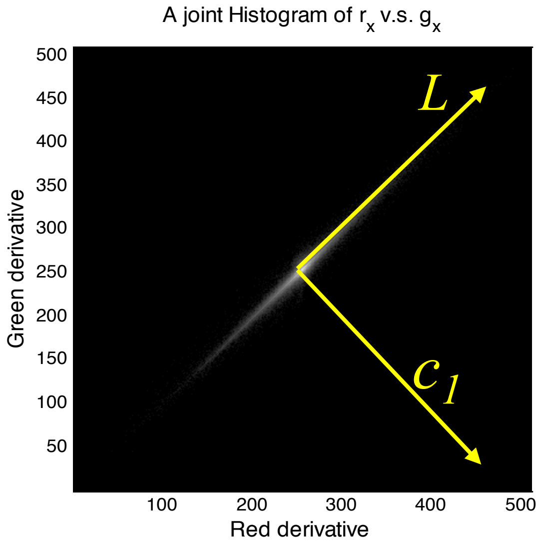
- The histogram shapes are highly stable over various images
- CC directions in 1x2 and 2x1 neighborhoods are the x-derivative and y-derivatives
- For kxk neighborhoods, all CC direction except one are high-frequency kernels (DC=0)
- The high-frequencies in different channels are highly probable to be identical.

Opponent Representation

- Let's consider the joint histogram of CC directions
- We define a new color basis (l,c_1,c_2):

$$\begin{pmatrix} l \\ c_1 \\ c_2 \end{pmatrix} = T \begin{pmatrix} R \\ G \\ B \end{pmatrix} \quad where \quad T = n \begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & 0 \\ 1 & 1 & -2 \end{pmatrix}$$

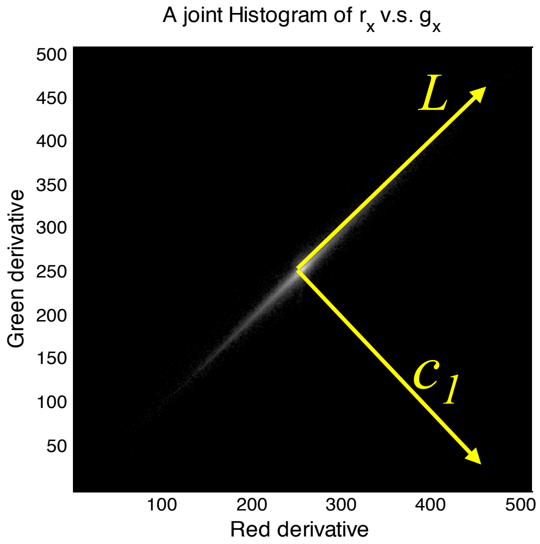
sum to 0



l – luminance
C_1- red/green
C_2 – blue/yellow

Observations:

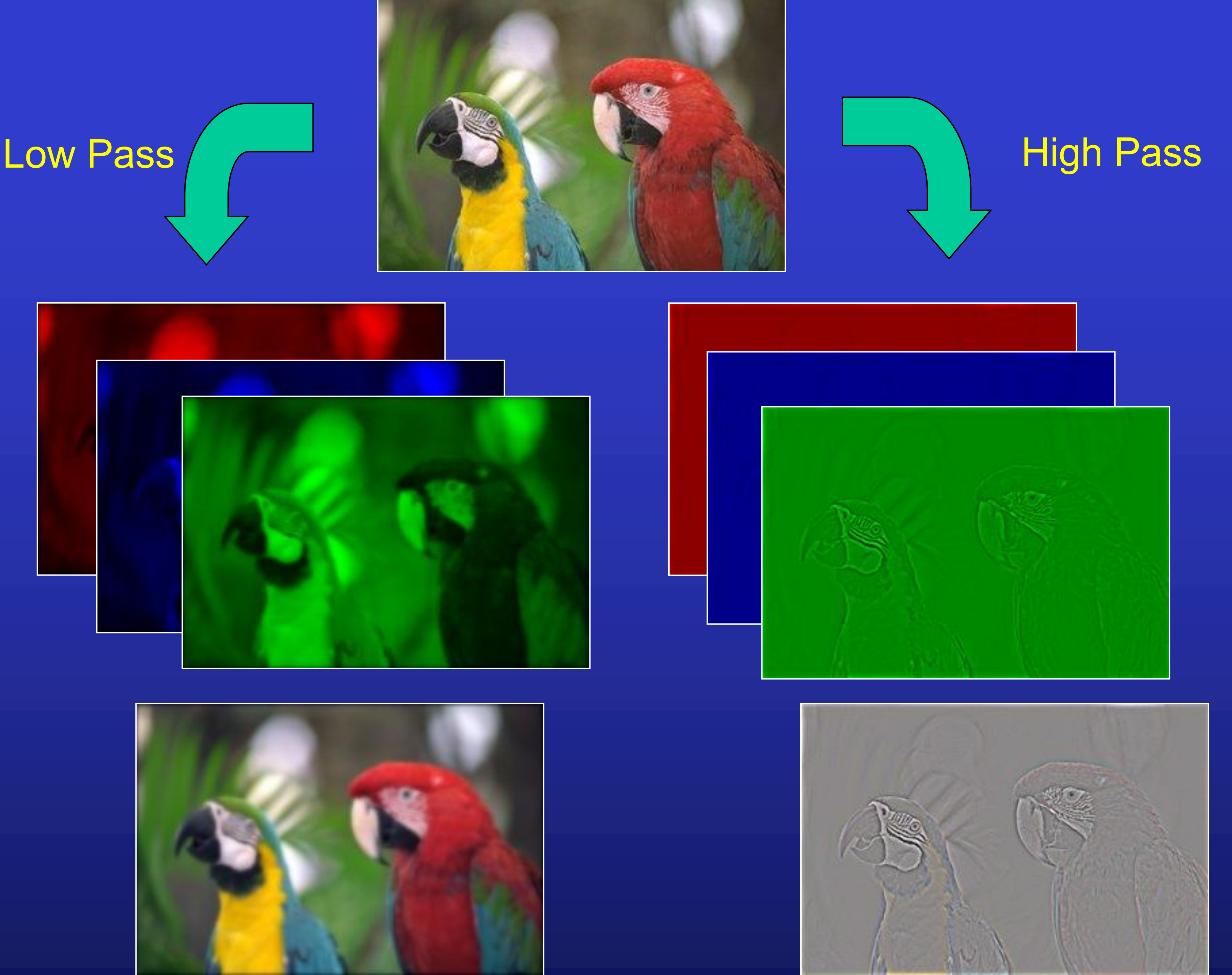
- The L channel encodes the **luminance**
- C_1 and C_2 channels encode the **chrominance**
- In the chrominance channels, high freq. are attenuated (canceled out)
- In the luminance channel, high freq. are maintained.
- The 3 opponent channels are uncorrelated in the high freq.



High derivatives

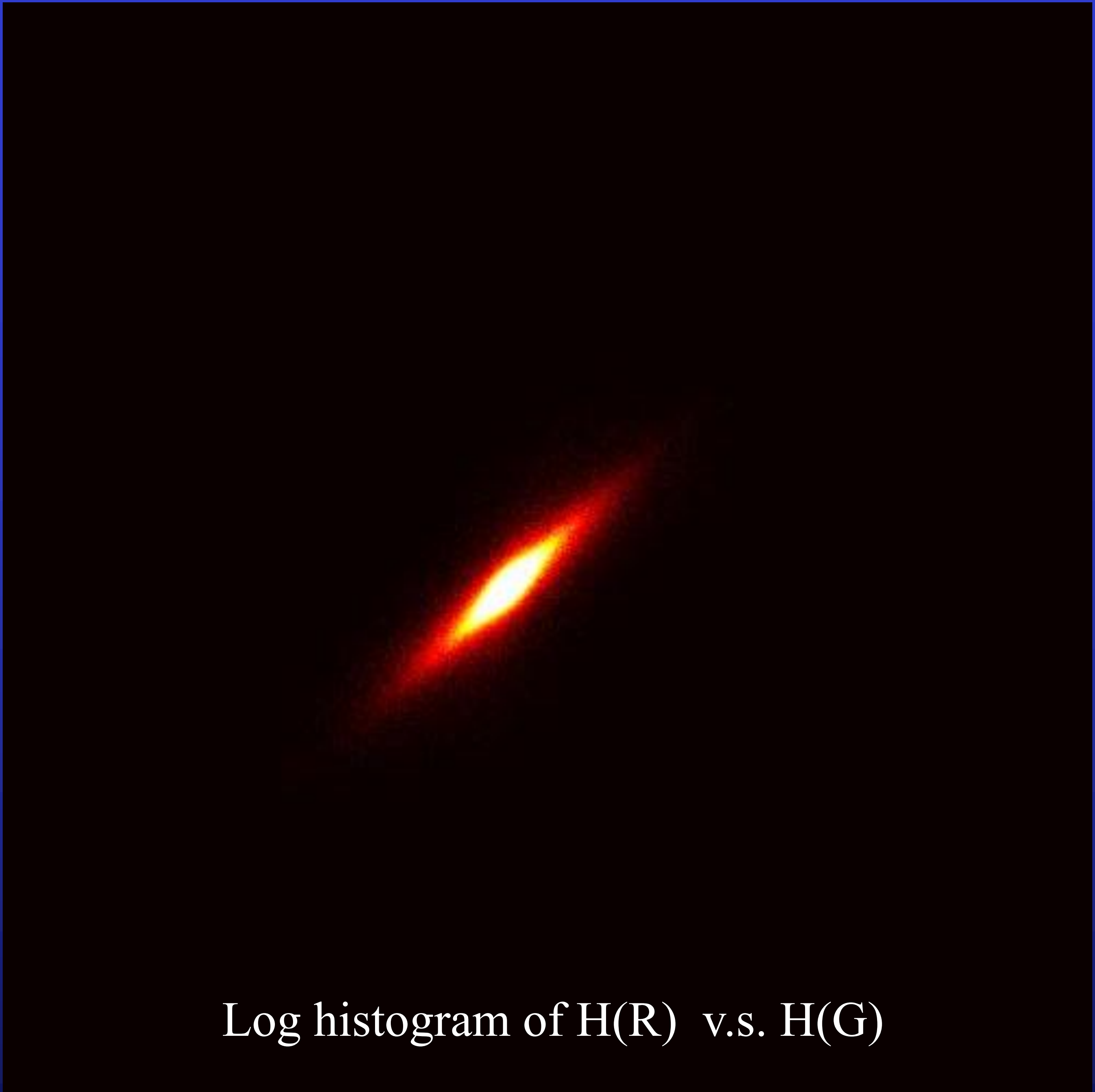
Low derivatives

Low derivatives

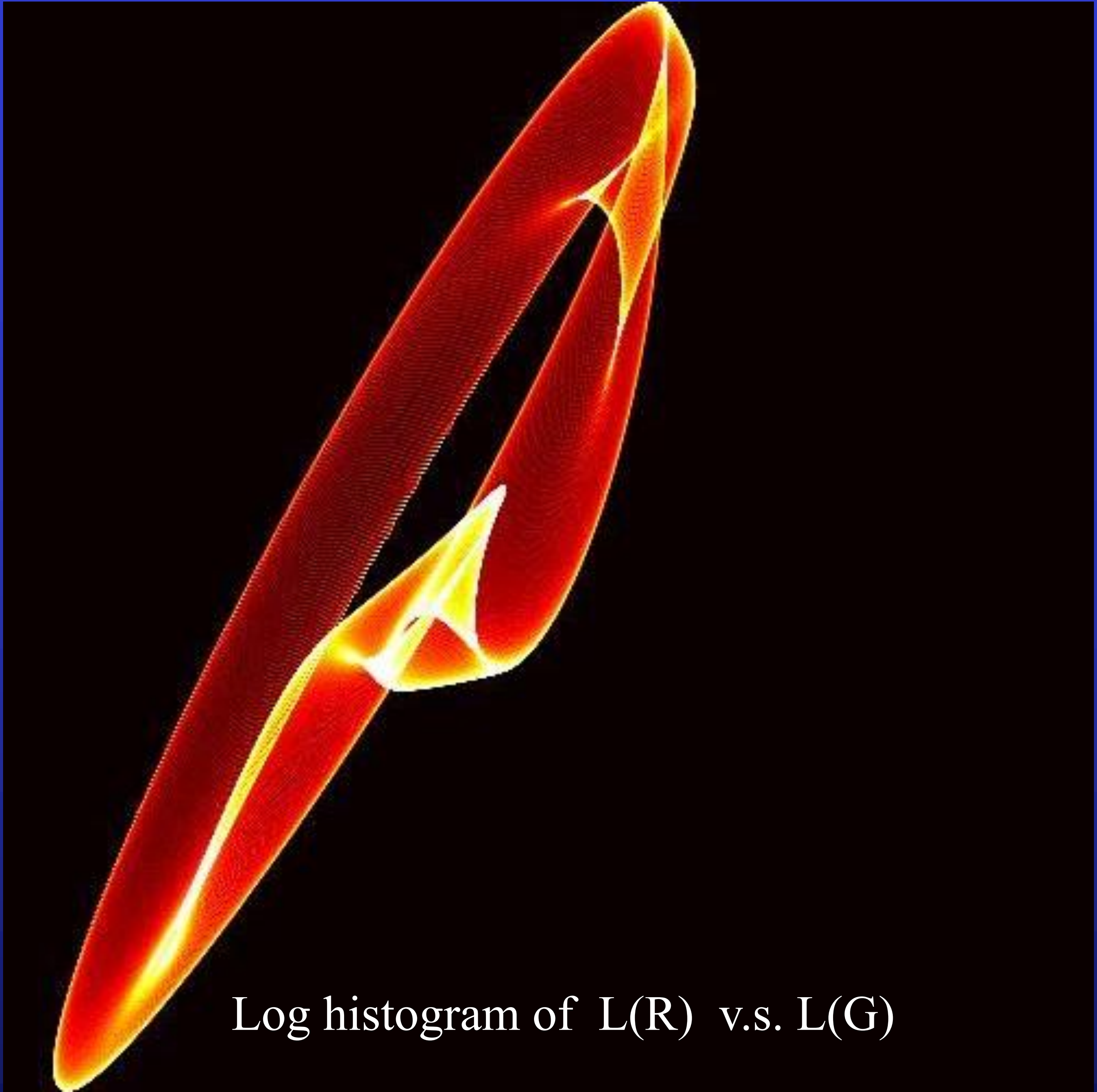


Low Pass

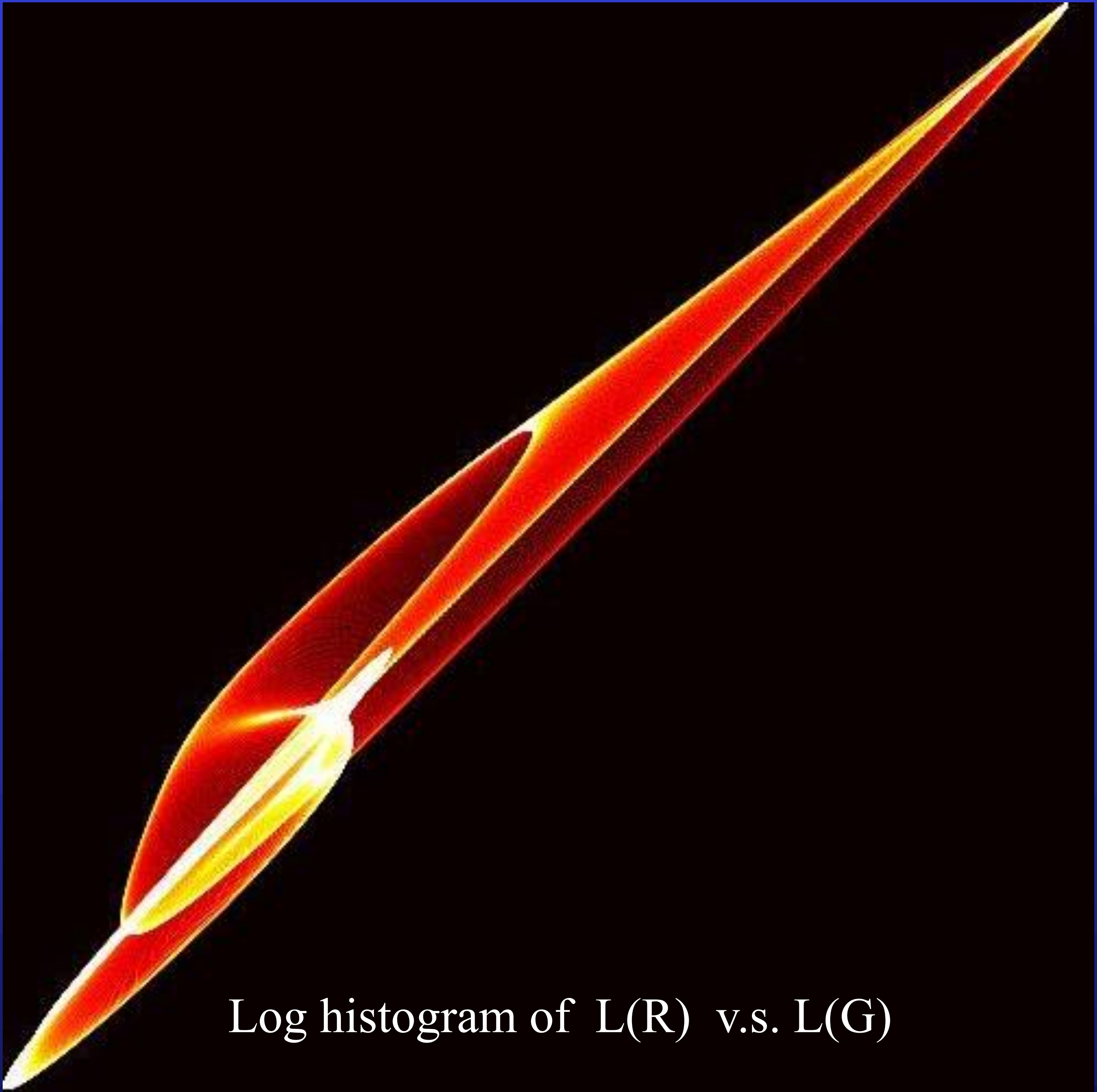
High Pass



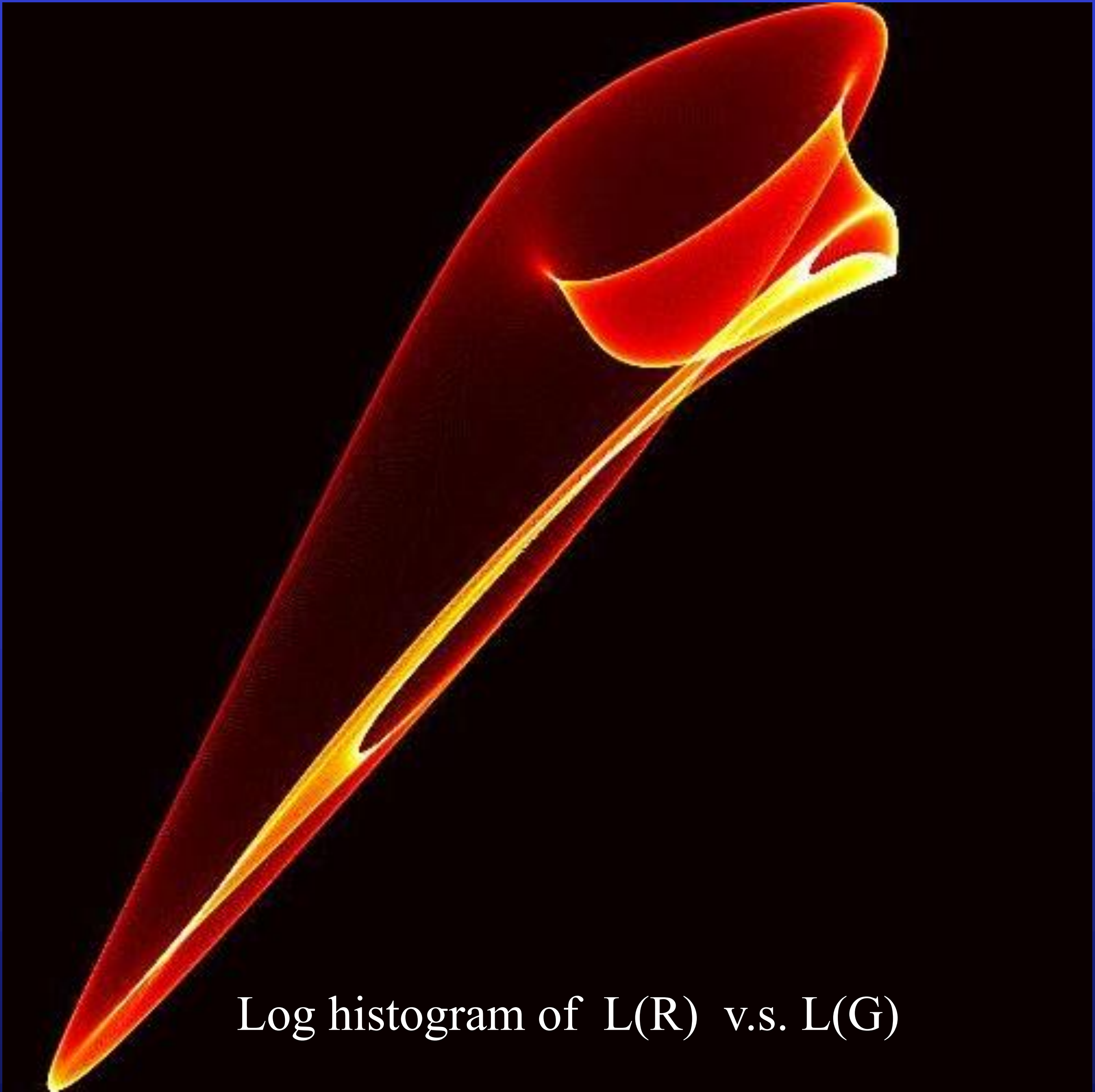
Log histogram of H(R) v.s. H(G)



Log histogram of L(R) v.s. L(G)



Log histogram of L(R) v.s. L(G)



Log histogram of L(R) v.s. L(G)

CCA Observations:

- Observation 1: Chrominance channels lack high spatial freq. because these are canceled out.
- Observation 2: The HVS is insensitive to high spatial freq. in the chrominance domain because it is improbable to have high freq. in those channels.
- Observation 3: Luminance/chrominance representation de-correlates the high spatial frequencies.

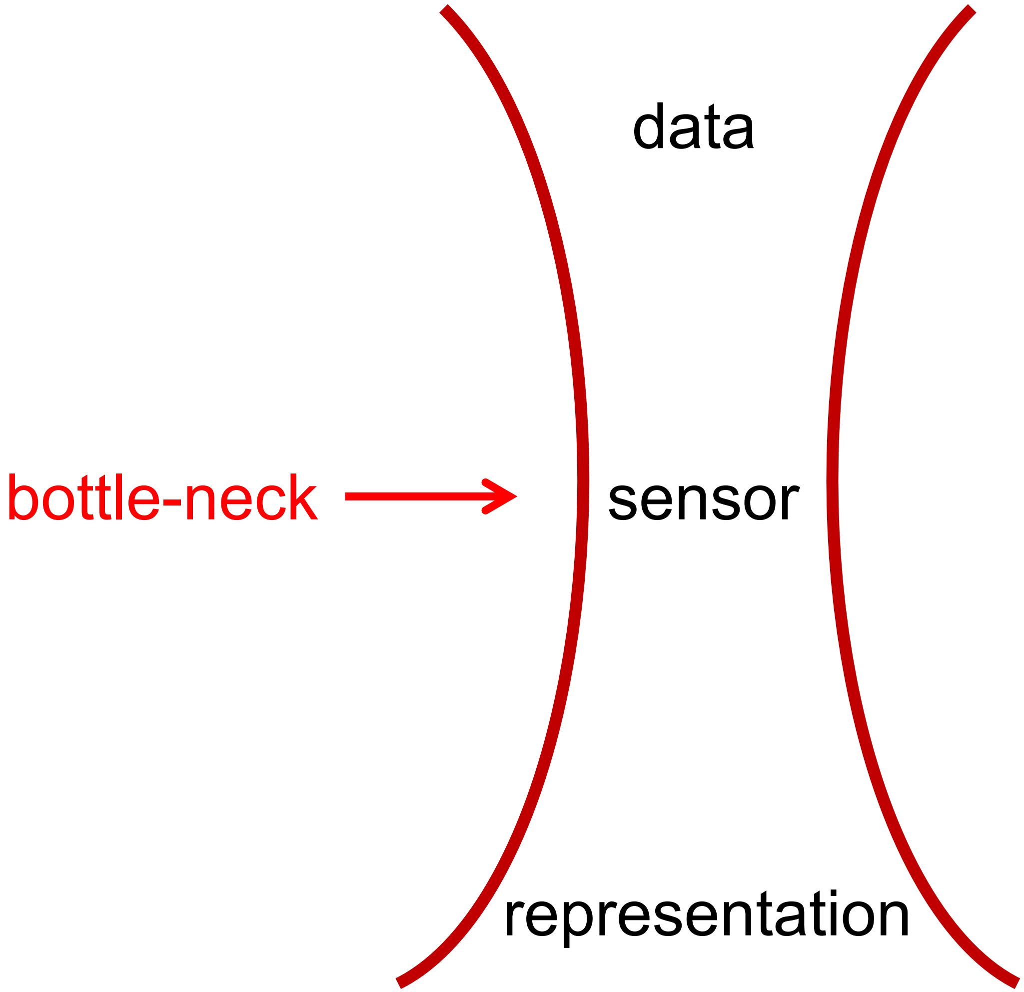
Efficient Reconstruction

Due to lack of space in the retina (fovea) we must plan spatio-chromatic sampling in an optimal way.

- Prop. 1: Non-uniform distribution of cones in the retina is a direct outcome of optimally sampling the spatio-chromatic space while taking into account the statistical prior.
- Prop. 2: Luminance/chrominance representation is a **necessary step** towards image reconstruction from the sensory data.

Lack of sensory data → Efficient Reconstruction

Efficient Reconstruction



data

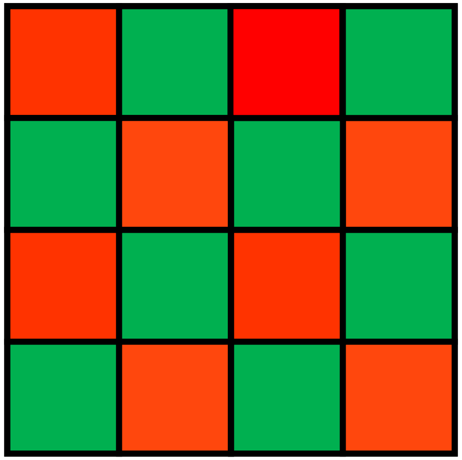
bottle-neck →

sensor

representation

Reconstruction in Opponent Space

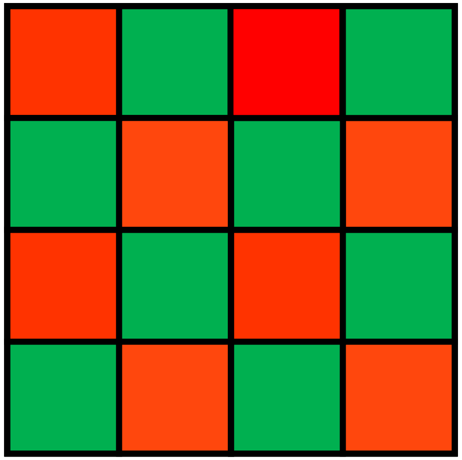
- For illustration, assume two types of sensors: R and G.
- What is the reconstructed resolution when equal-rate of sampling is applied?



- Nyquist already answered this!

Reconstruction in Opponent Space

- Let’s assign more Red sensors.
- What is the reconstructed resolution of the Red channel?
- Nyquist already answered this!
- What is the reconstructed resolution of the Green channel?
- The resolution of the Red channel.



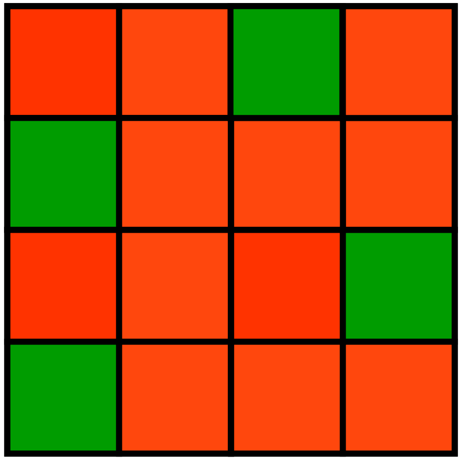
Reconstruction in Opponent Space

Dense sampling $R = R_H + R_L$

Sparse sampling $G = G_H + G_L$

$$\hat{R} = interp(R)$$

$$\hat{G} = interp(G - R) + R$$



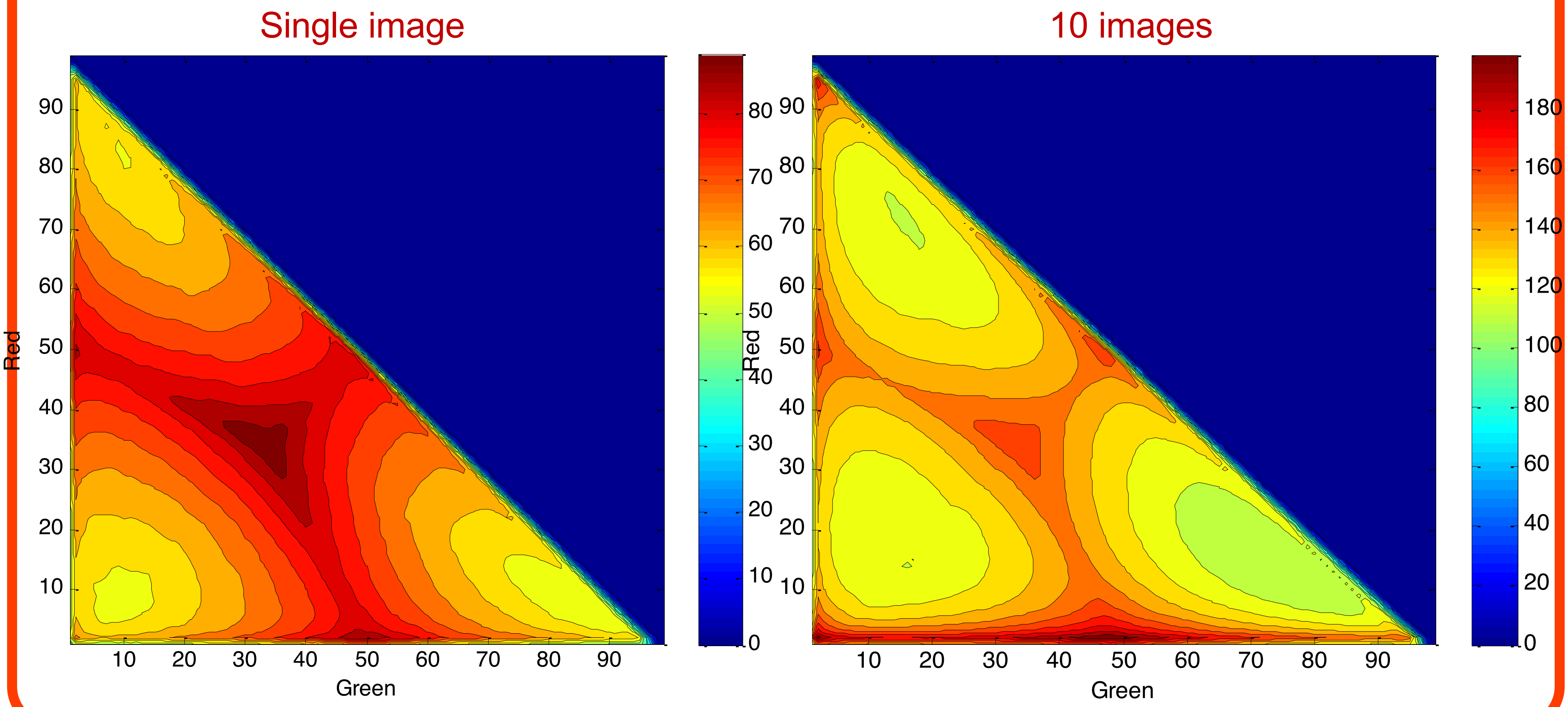
🟥	🟥	🟩	🟥
🟩	🟥	🟥	🟥
🟥	🟥	🟥	🟩
🟩	🟥	🟥	🟥

||

Sparse sampling is sufficient $G_L - R_L$ Opponent Channel

Let’s try it on real images:

- Sample images using RGB sampling rates with probability $\alpha, \beta, (1 - \alpha - \beta)$
- Apply simple reconstruction using the suggested scheme



To Conclude:

Efficient Reconstruction is proposed as a model for retinoscopic representation

- Q1: Why is the distribution of L M S cones non uniform?
- A1: Optimal sampling
- Q2: Why does the HVS encode color information in opponent space?
- A2: Data reconstruction
- Q3: Why is the HVS less sensitive to high spatial frequencies in the chrominance channels?
- A3: High freq. in chrominance channels are attenuated. Sparsely sampled chrominance channels is sufficient.

Summary:

- We propose a new viewpoint which unifies several cortical observations.
- Unlike the current belief that the cortical representation stems from “efficient coding”, we argue that the sensory setup as well as the cortical representation are outcome of “efficient reconstruction”.
- The ultimate goal is to maximize the amount of information (resolution) extracted from a limited amount of samples.


THE
END

Questions:

- Can we design a psychophysical/physiological experiments verifying the propositions?
- Does the LMS sampling rate conform with the observed RGB correlation values?

CCA V.S. PCA

	CCA	PCA
Variables	distinct entities	augmented
Mutual Correlations	Maximizes	Minimizes
Affine Trans.	independent	dependent
Consideration	*between* classes	*between* and *within* classes