MANIFOLD LEARNING IN VISION

NON-LINEAR REPRESENTATIONS FOR TEXTURE AND COLOR CHARACTERIZATION

JESÚS MALO

DEPT. OPTICS
IMAGE PROCESSING LAB.
UNIVERSITAT DE VALÈNCIA
(SPAIN)

GUANAJUATO 2011

SUMMARY : MANIFOLD LEARNING IN VISION

① Something about me : research agenda THE BARLOW HYPOTHESIS

② Statistical features of images and colors

③ Phenomenology of human vision (in textures and color)

④ Two unsupervised learning algorithms SPCA and RBIG

⑤ Neuroscience from statistics and viceversa

⑥ Applications in image processing

- Coding
- Denoising
- Classification
- Synthesis
- Color constancy
- Image quality

$$\text{SURVIVAL} \Rightarrow \text{OPTIMAL} \left\{ \begin{array}{l} \text{ANALYSIS} \\ \text{CHOICES} \end{array} \right.$$

FEATURES ?

SENSORS ?

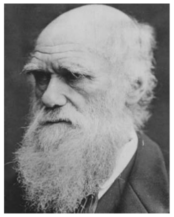
RESOLUTION ?

PDF of EVENTS ?

BAYES

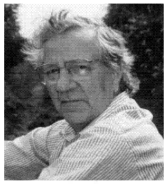
1702 - 1761

DARWIN



1809 - 1882

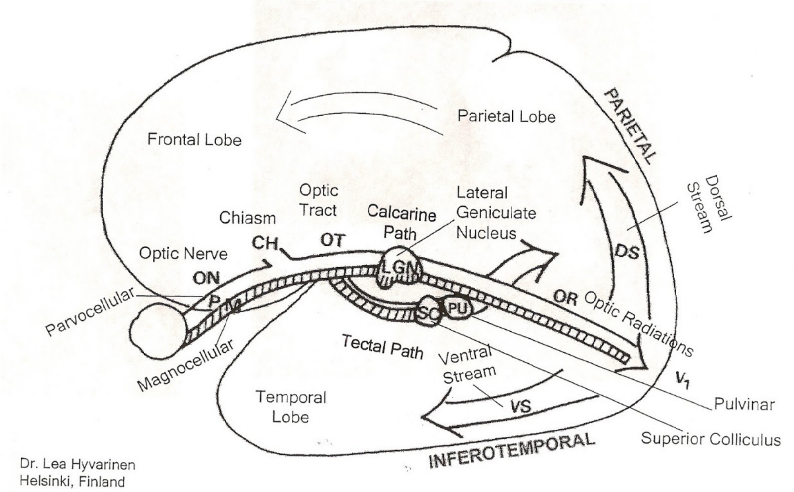
BARLOW



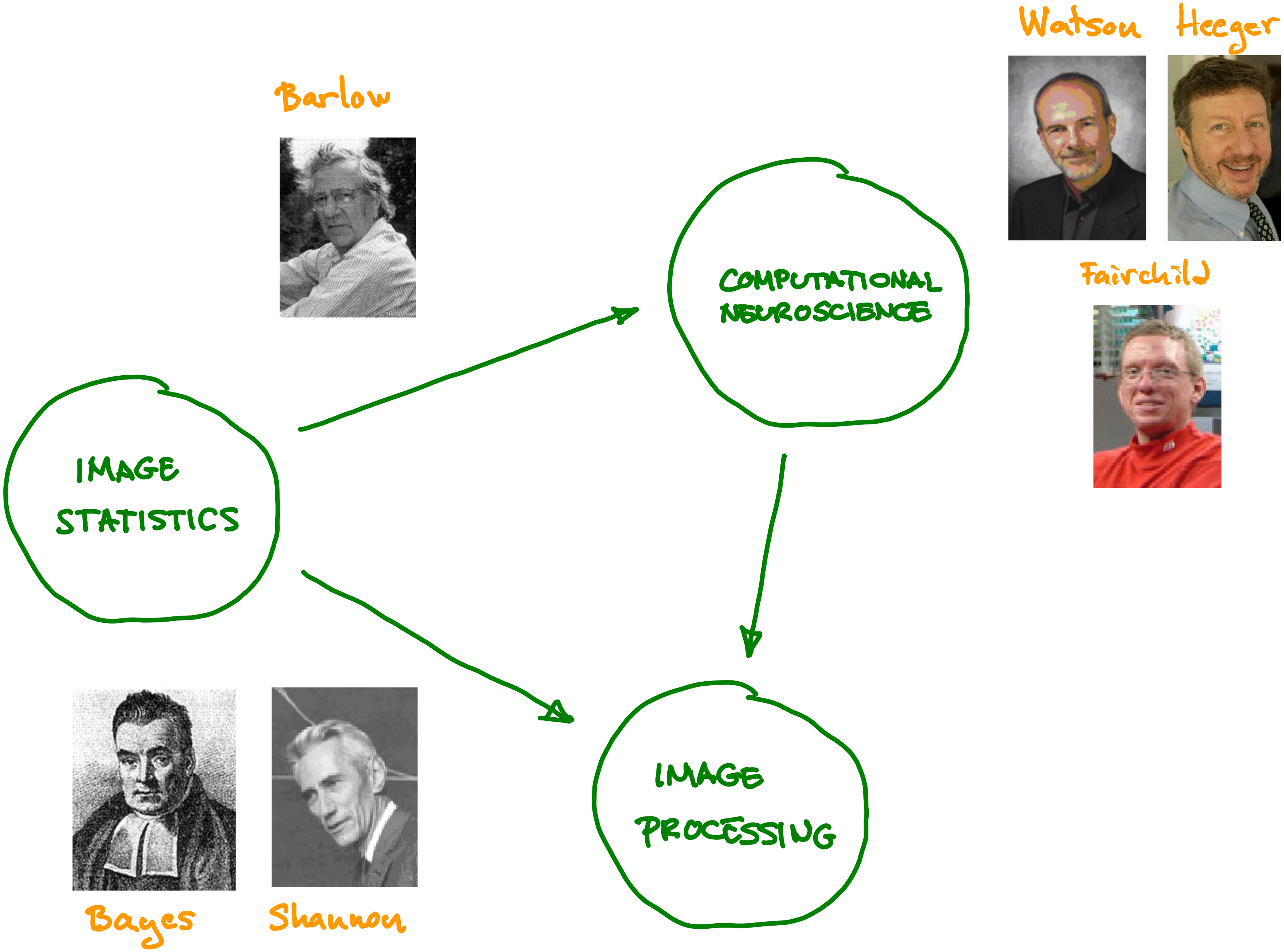
1921 –

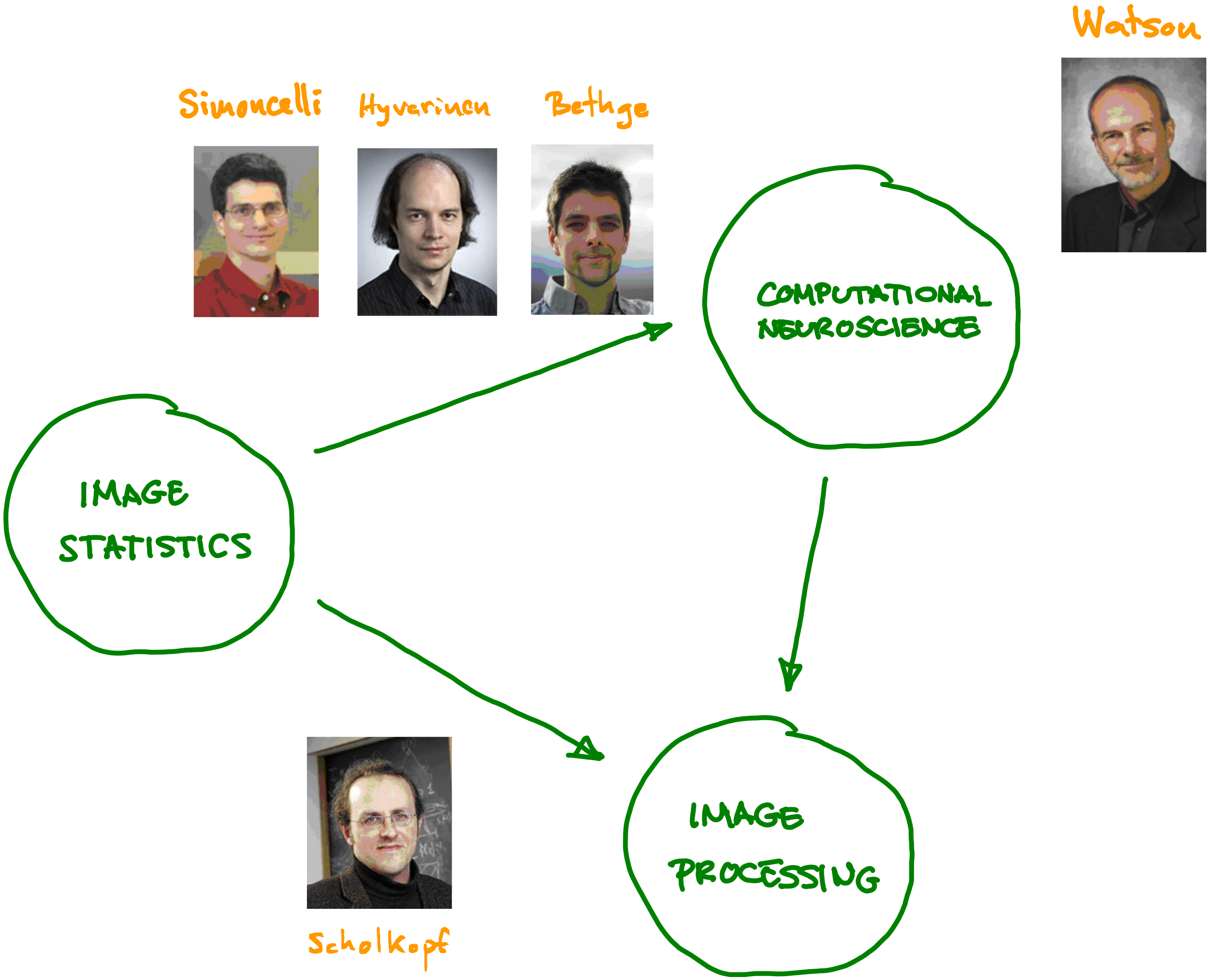
Diagram 3

Pathways for Visual Information in the Brain

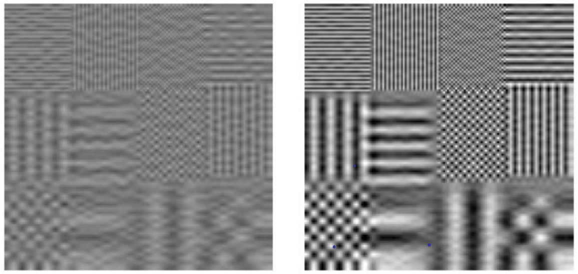


Dr. Lea Hyvarinen
Helsinki, Finland

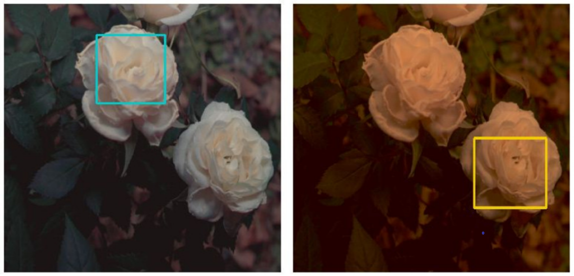
Watson
Heeger
Barlow
COMPUTATIONAL NEUROSCIENCE
Fairchild
IMAGE STATISTICS
IMAGE PROCESSING
Bayes
Shannon


Watson
Simoncelli
Hyvarinen
Bethge
COMPUTATIONAL NEUROSCIENCE
IMAGE STATISTICS
IMAGE PROCESSING
Scholkopf

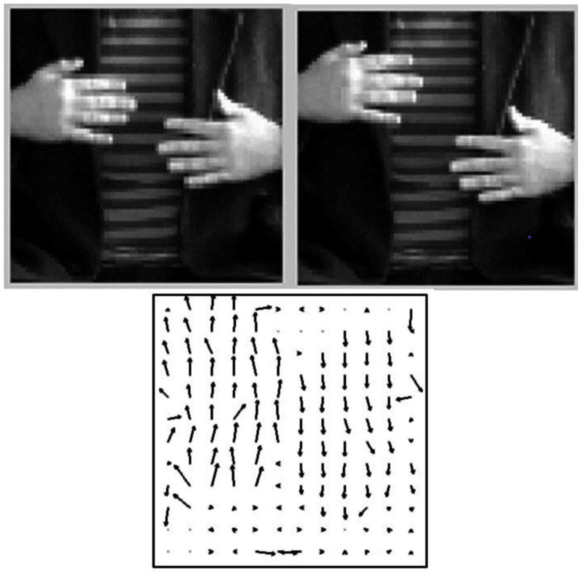
TEXTURE



COLOR



MOTION



COMPUTATIONAL NEUROSCIENCE

Texture:
- Gabor receptive fields
- Linear Contrast sensitivity
- Non-linear Contrast Adaptation

Color:
- Opponent channels
- Color discrimination
- Color constancy

Motion:
- Optical flow from 3D Gabor R.F.
- Motion adaptation divisive normaliz.

Statistics:
- Information theory & ICA

RESEARCH AGENDA: We are actually doing it!

0 Non-linearities and adaptation of color vision from Sequential Principal Curves Analysis. Subm. Neur. Comp

1 Complex ICA of natural images. ICANN 2011

2 V1 model computes a sort of non-linear ICA. Neur. Comp. 2010

3 V1 model for image distortion assessment. JOSA A. 2010

4 PDF of natural images using Rotation-Based Iterative Gaussianization. IEEE Neural Networks 2011

5 Image denoising using Kernels based on natural image statistics J. Machine Learning. Res. 2010

6 V1 non-linearities emerge from non-linear ICA. Network 2006

7 Divisive normalization for efficient perceptual coding. IEEE Im. Proc. 2006

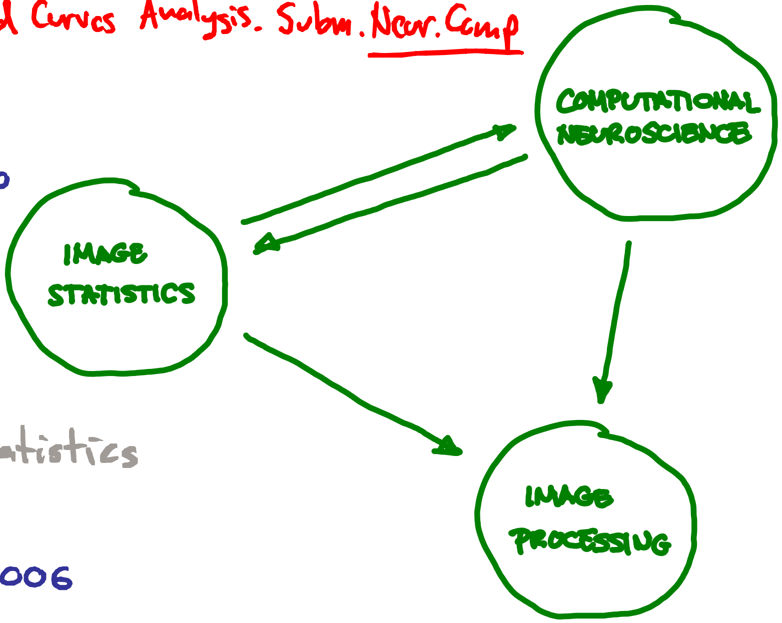
8 Divisive normalization and SVM for transform coding. J. Mach. Learn. Res. 2007

9 Contrast Sensitivity function-based SVM for image coding. IEEE Neur. Networks 2005

10 Corresponding pair procedure to simulate dichromatic color vision JOSA A. 2004

11 Perceptual motion estimation for video coding. IEEE Im. Proc. 2001

12 Perceptual optical flow for motion segmentation in MPEG4. Electr. Letts. 2000



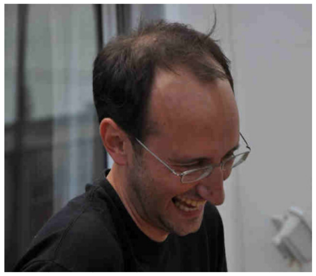
I am not alone!: IMAGE PROCESSING LAB (UNIVERSITAT DE VALENCIA)

http://isp.uv.es

http://www.uv.es/vista/vistavalencia

CURRENT


Valero
LAPARRA


Gustavo
CAMPS-VALLS


Jordi
MUÑOZ

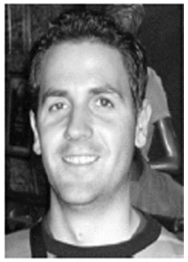

Sandra
JIMENEZ


Vicent
TALENS

PAST


Juan
GUTIERREZ
(Comp. Sci
U.V)


Irene
EPIFANIO
(Maths
UJI)


Gabriel
GOMEZ
(Analog
Devices)

SUMMARY : MANIFOLD LEARNING IN VISION

1. Something about me : research agenda THE BARLOW HYPOTHESIS
2. Statistical features of images and colors
3. Phenomenology of human vision (in textures and color)
4. Two unsupervised learning algorithms SPCA and RBIG
5. Neuroscience from statistics and viceversa
6. Applications in image processing
    - Coding
    - Denoising
    - Classification
    - Synthesis
    - Color constancy
    - Image quality

SUMMARY : MANIFOLD LEARNING IN VISION

① Something about me: research agenda THE BARLOW HYPOTHESIS

② **Statistical features of images and colors**

③ Phenomenology of human vision (in textures and color)

④ Two unsupervised learning algorithms SPCA and RBIG

⑤ Neuroscience from statistics and viceversa

⑥ Applications in image processing

- Coding
- Denoising
- Classification
- Synthesis
- Color constancy
- Image quality

② BASIC STATISTICAL FEATURES OF IMAGES AND COLORS

2.1 Images and colors as vectors

2.2 Unsupervised learning

2.3 Properties of natural image patches

- Second order statistics (linear) PCA and DCT
- Higher order statistics linear ICA and Wavelets
- What linear representations cannot capture

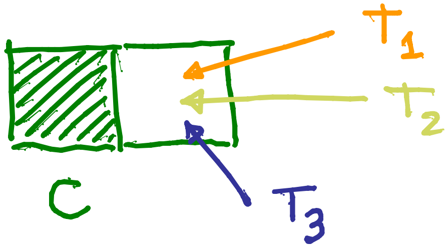
2.4 Properties of natural colors

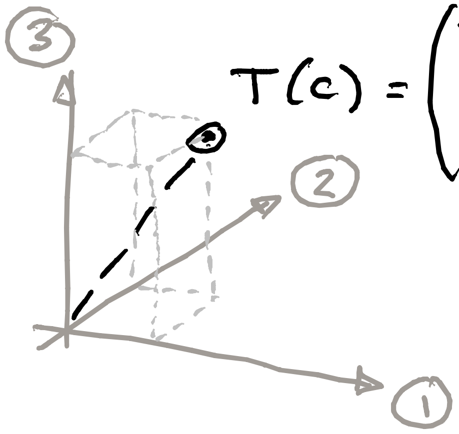
- Second order statistics (in RGB-like representations)
- PCA and opponent channels
- Linear and non-linear changes in color manifolds

② BASIC STATISTICAL FEATURES OF IMAGES AND COLORS

2.1 Images and colors as vectors

COLOR: in classical tristimulus colorimetry, color of a small region is a 3D vector

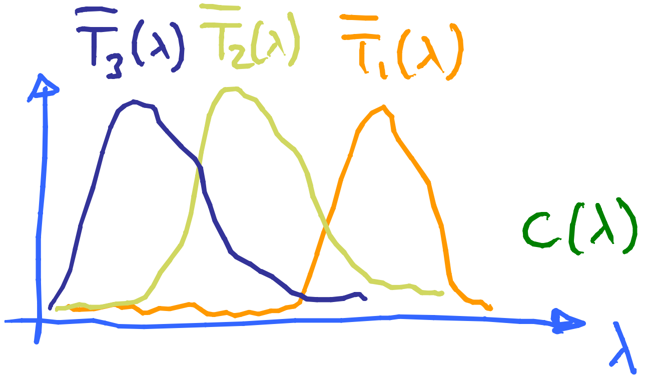




$$T(c) = \begin{pmatrix} T_1(c) \\ T_2(c) \\ T_3(c) \end{pmatrix}$$

Grassmann laws [Stiles 82]

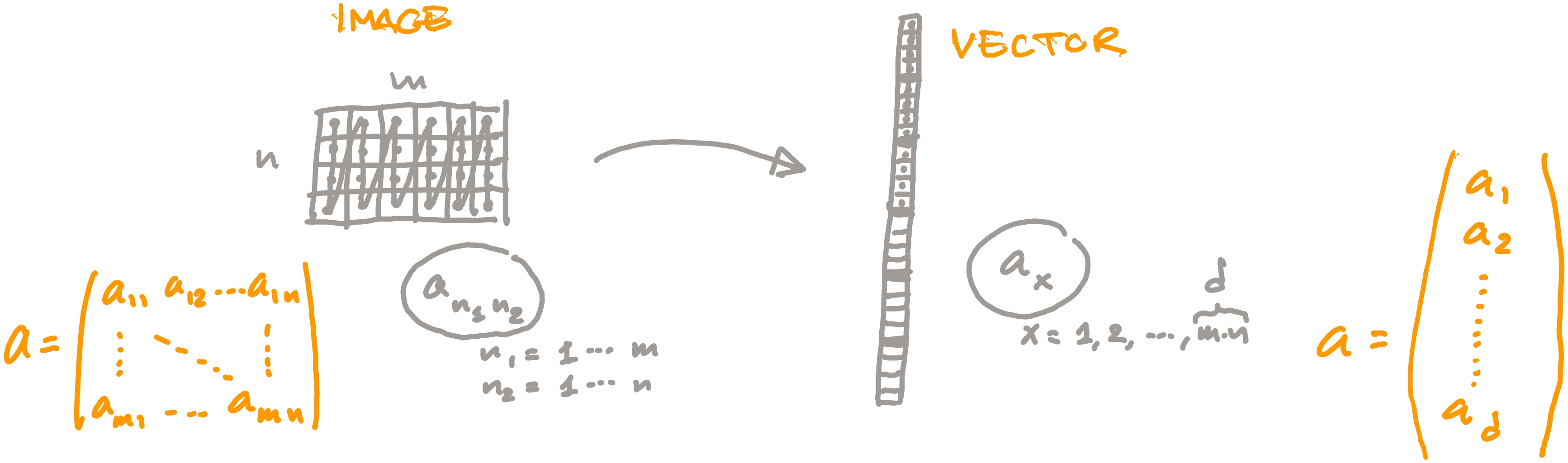
$$T_i(c) = \int c(\lambda)\,\bar{T}_i(\lambda)\,d\lambda$$



② BASIC STATISTICAL FEATURES OF IMAGES AND COLORS

2.1 Images and colors as vectors

IMAGES: $n \times m$ images can be considered as $d \times 1$ vectors (with $d = n \cdot m$)



$$a = \begin{bmatrix} a_{11} & a_{12} \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}$$

$\boxed{a_{n_1 n_2}}$

$n_1 = 1 \cdots m$

$n_2 = 1 \cdots n$

$\boxed{a_x}$

$x = 1, 2, \cdots, \overbrace{m \cdot n}^{d}$

$$a = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_d \end{pmatrix}$$

2D MATRIX $\longrightarrow$ $\overbrace{m \cdot n}^{d}$ D VECTOR

Problem: $\vec{n}$ has strict physical meaning (position) x has spatial meaning (but not obvious)

Benefit: You apply linear algebra.

② BASIC STATISTICAL FEATURES OF IMAGES AND COLORS

2.1 Images and colors as vectors

* Images and colors are {
d-dimensional vectors (d is HUGE!)
3-dimensional vectors

* This can be extended to color images and sequences...

FUNDAMENTAL QUESTIONS

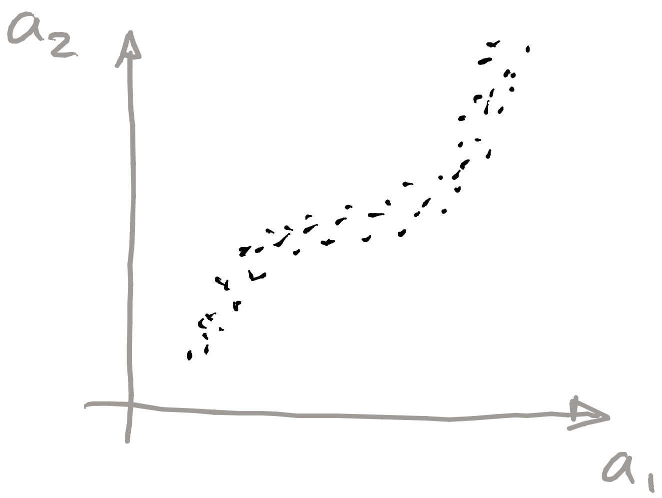
- Where do these vectors live?
- Do they fill the space?
- Does its distribution have any structure at all?

② BASIC STATISTICAL FEATURES OF IMAGES AND COLORS

2.2 A note on unsupervised learning

UNSUPERVISED LEARNING: techniques that automatically find structure of multidimensional samples

- clusters
- low dimensional subspaces
- fundamental directions or symmetries



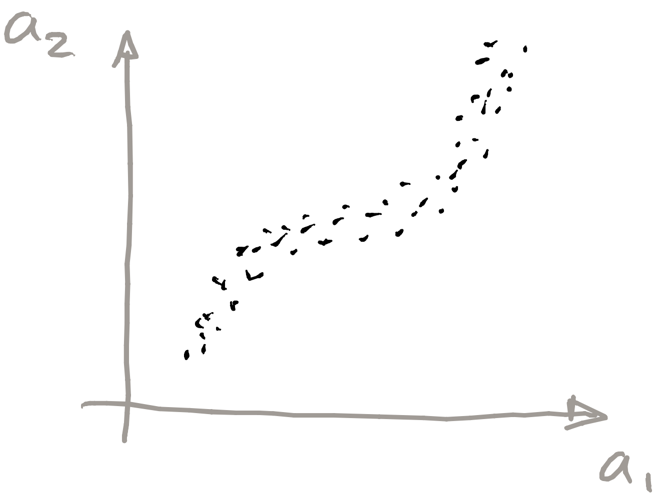
② BASIC STATISTICAL FEATURES OF IMAGES AND COLORS

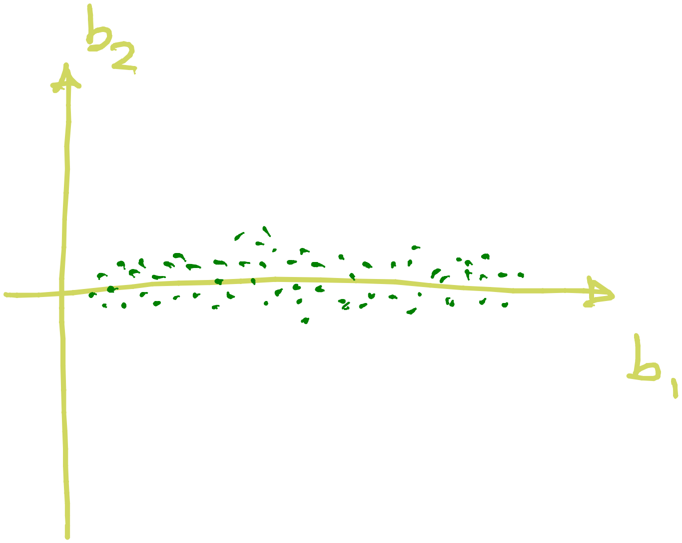
2.2 A note on unsupervised learning

WHAT IS STRUCTURE ?:

Multiinformation [Cover & Thomas 92]: $I(a_1, \dots, a_n) = \sum_{x=1}^{d} H(a_x) - H(a)$

Information that variable a_i give us about variables a_j





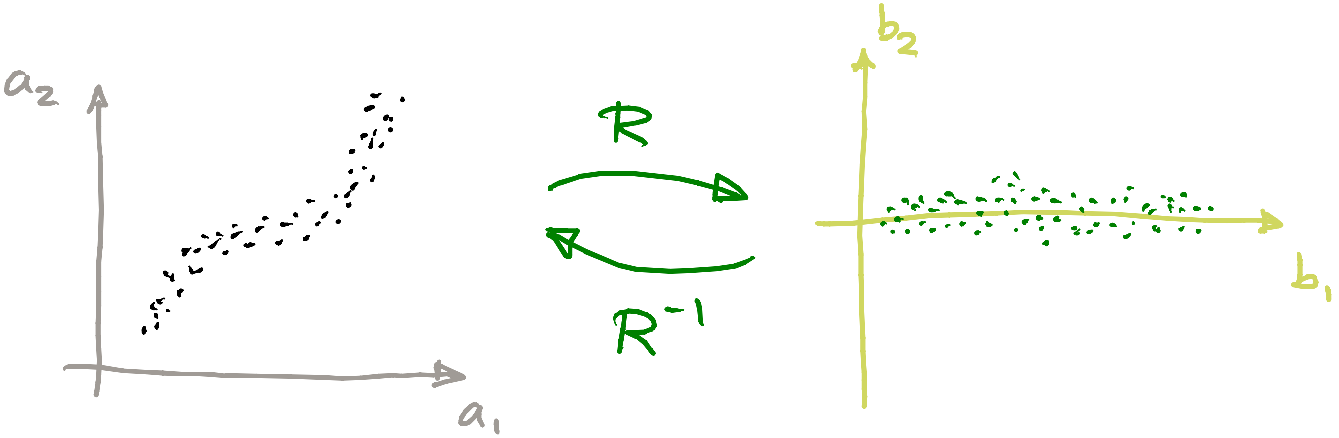
② BASIC STATISTICAL FEATURES OF IMAGES AND COLORS

2.2 A note on unsupervised learning

SINCE A possible definition of STRUCTURE is MULTIINFORMATION

⇓

Describing the structure is finding transforms that remove I



If $R : a \longrightarrow b$ removes I, the structure of a is encoded in R

Feature extraction transforms with such a property are called ICA $\left\{ \begin{array}{l} \text{- LINEAR} \\ \text{- NON-LINEAR} \end{array} \right.$

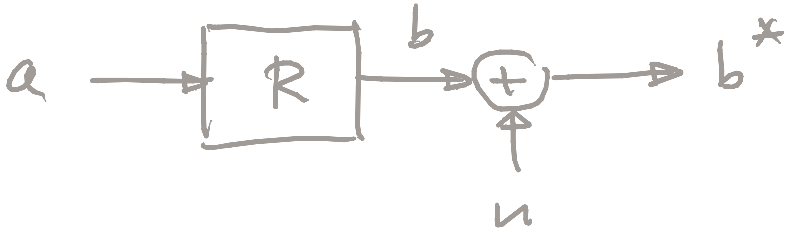
[Comon94, Hyvarinen00]

② BASIC STATISTICAL FEATURES OF IMAGES AND COLORS

2.2 A note on unsupervised learning

ANOTHER WAY TO SEE IT: Optimal sensory system design (BARLOW !)

Infomax [Linsker 82, Bell & Sejnowski 95]



Infomax: maximize the information transference $I(a,b^*)$

$$I(a,b^*) = H(b^*) - H(b^* \mid a) = H(b^*) - H(n)$$

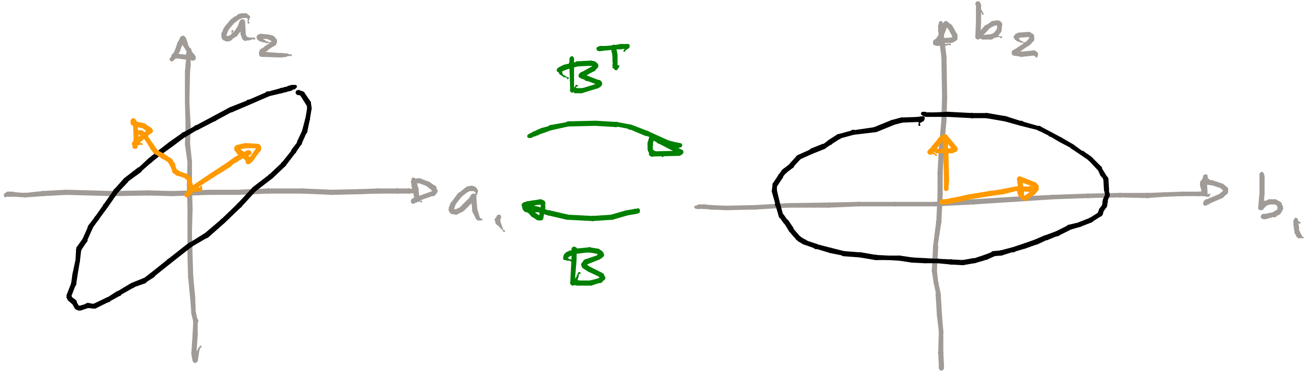
$$\Downarrow$$

Optimal R is the one that maximizes $H(b) \Rightarrow$ ICA

② BASIC STATISTICAL FEATURES OF IMAGES AND COLORS

2.2 A note on unsupervised learning

A well know example for Gaussian sources : PCA



$$\Sigma_b = B^T \cdot \Sigma_a \cdot B \qquad (B \text{ diagonalizes } \Sigma_b)$$

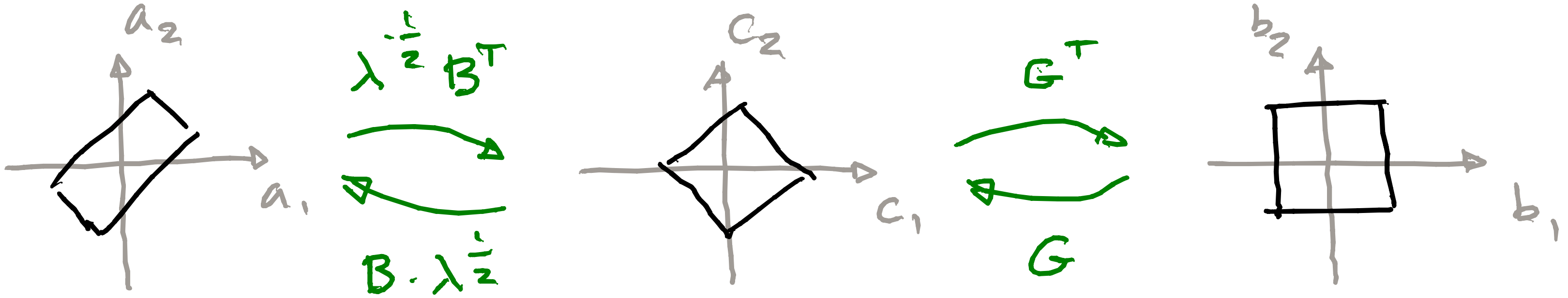
$$b = B^T \cdot a$$

Structure is encoded in the eigenvectors of Σ_a (colums of B)

② BASIC STATISTICAL FEATURES OF IMAGES AND COLORS

2.2 A note on unsupervised learning

A more recent (more general) example: linear ICA [Comon 94]



Here structure is in the rows of W and colums of A

$$b = \underbrace{G^T \cdot \lambda^{-\frac{1}{2}} B^T}_{} \cdot a = W \cdot a$$

$$a = \underbrace{B \cdot \lambda^{\frac{1}{2}} \cdot G}_{} \cdot b = A \cdot b$$

where $\begin{cases} W \equiv \text{Separating matrix} \\ A \equiv \text{Mixing matrix} \end{cases}$

$A = W^{-1}$

where $\begin{cases} * \ \lambda^{-\frac{1}{2}} B^T \equiv \text{Whitening} \\ * \ G \equiv \text{Rotation that maximizes marginal non-Gaussianity} \end{cases}$

Demo:

[Cardoso 03] $I = C + J - J_m$ $\begin{cases} \cdot\ C \equiv \text{2nd order correlation} \\ \cdot\ J \equiv \text{Global non-Gauss.} \\ \cdot\ J_m \equiv \text{Marginal non-Gauss.} \end{cases}$

② BASIC STATISTICAL FEATURES OF IMAGES AND COLORS

2.2 A note on unsupervised learning

A list of feature extraction / unsupervised learning techniques [Verleysen 07]

PERSONAL OPINION !

Linear
- PCA
- ICA

Non linear
- NL-PCA / NL-ICA
- Gaussianization - Projection Pursuit
- SOM - Vector Quantiz.
- Kernel generalizations of linear techniques KPCA, KICA
- Spectral methods
  - ISOMAP
  - LLE

GOOD: they charact. the PDF

~

BAD: they do not charact. the PDF or they are not intuitive

② BASIC STATISTICAL FEATURES OF IMAGES AND COLORS

FINE

* Visual information
- images
- colors
- sequences

can be seen as vectors

* There are a bunch of unsupervised learning techniques for optimal sensor (representation) design.

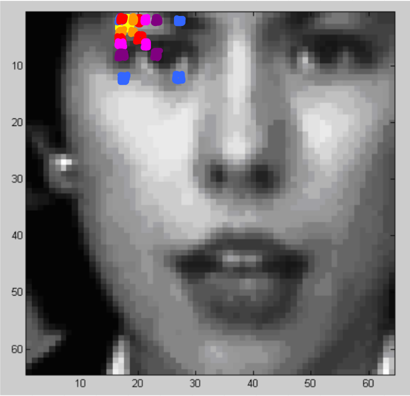
THE RELEVANT QUESTIONS:

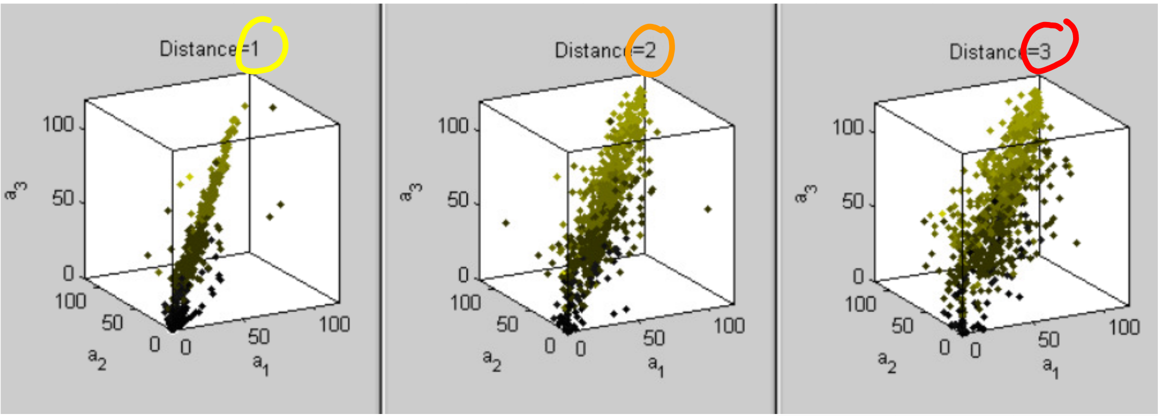
- Where do these vectors live?
- Do they fill the space?
- Does its distribution have any structure at all?
- How these distributions change with changes in obs. cond,

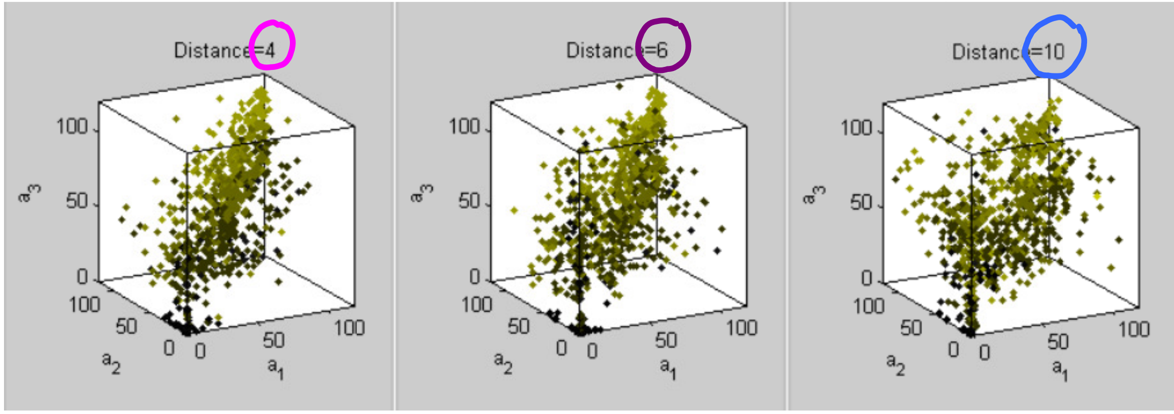
② BASIC STATISTICAL FEATURES OF IMAGES AND COLORS

2.3 Properties of natural image patches

Images are smooth!: relations
between neighbors



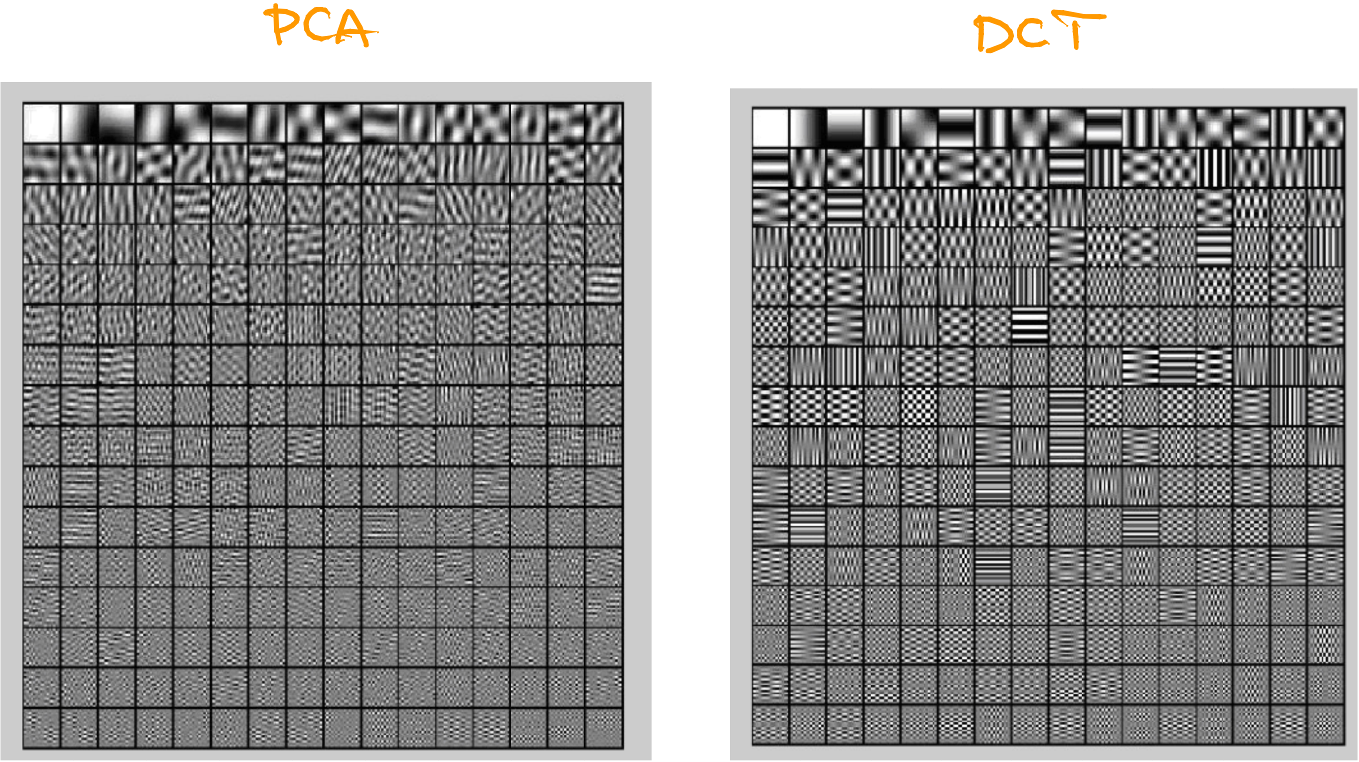




② BASIC STATISTICAL FEATURES OF IMAGES AND COLORS

2.3 Properties of natural image patches

Images are smooth ⇒ PCA ~ DCT (Clarke 81, Hankock 92)
(2nd order approach)



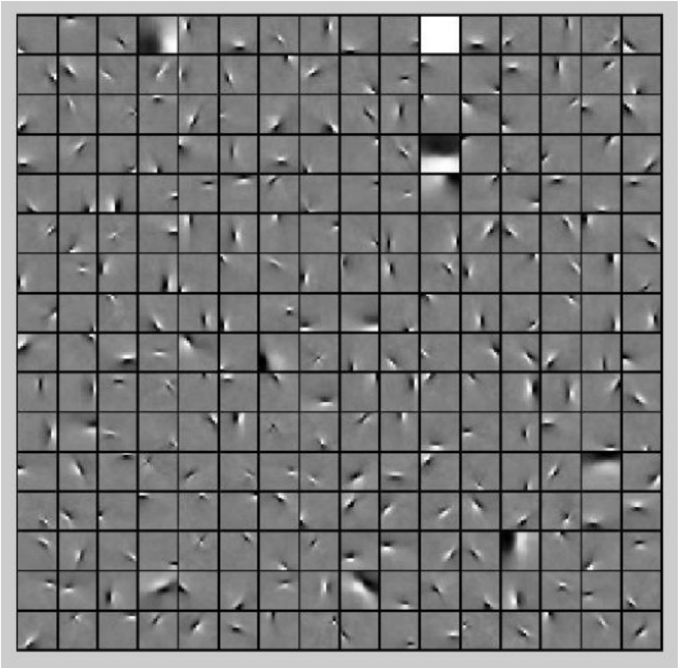
② BASIC STATISTICAL FEATURES OF IMAGES AND COLORS

2.3 Properties of natural image patches

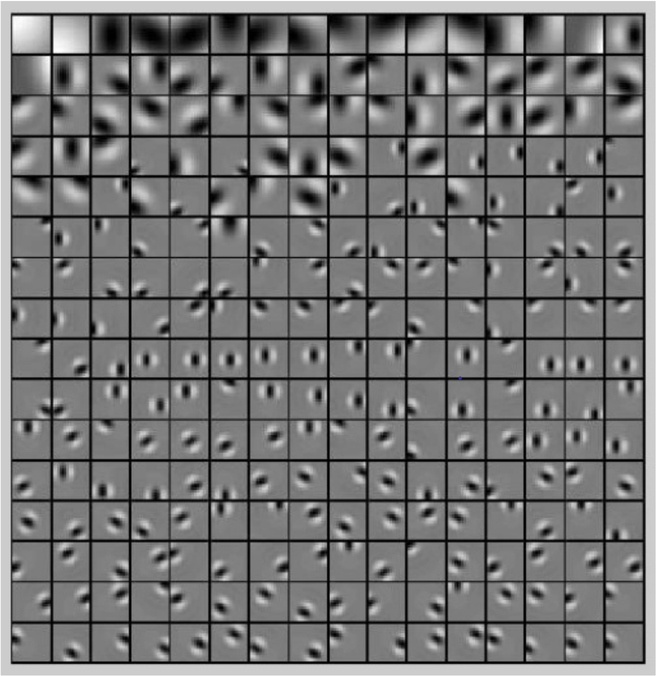
Images have distinct features (as edges) $\Rightarrow$ linear ICA ~ Wavelets
(Higher order – linear – approach)

[Olshausen&Field96]

linear ICA



Wavelets



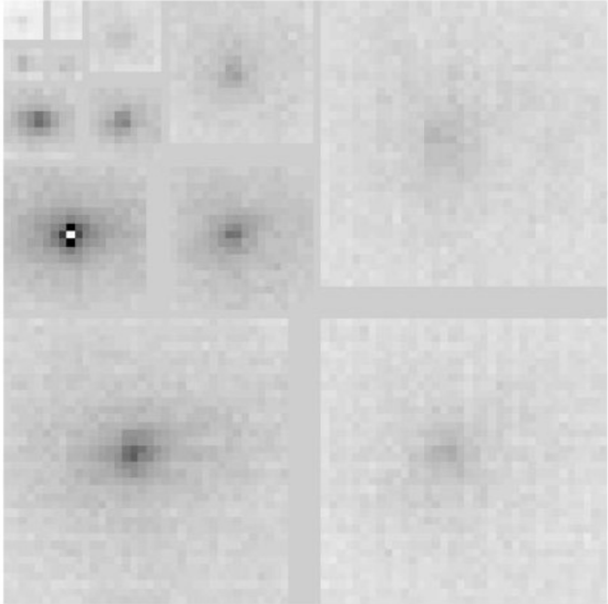
② BASIC STATISTICAL FEATURES OF IMAGES AND COLORS

2.3 Properties of natural image patches

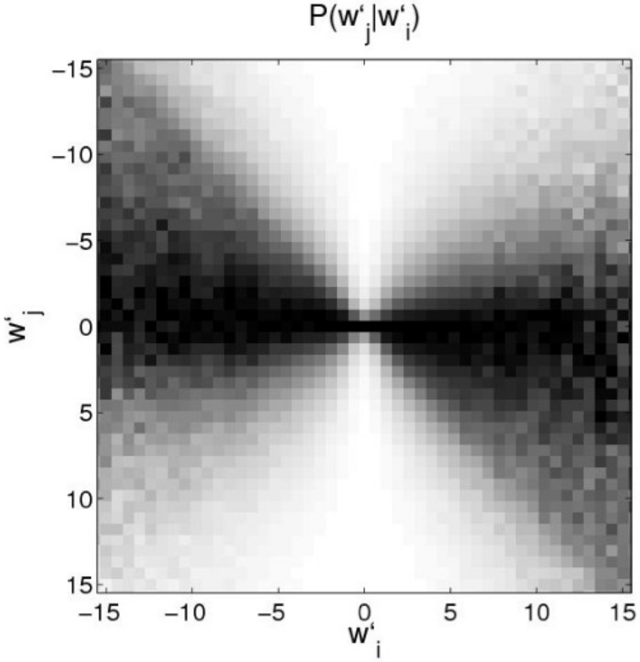
HOWEVER, images are more like that: $a = A \cdot b$

What linear transforms cannot capture!

[Simoncelli 97, Liu01, Hyvarinen03, Malo&Laparra 10]



$MI_{\mathbf{w}}$: 0.052, $[10^{-3}, 0.27]$ bits



② BASIC STATISTICAL FEATURES OF IMAGES AND COLORS

2.3 Properties of natural image patches

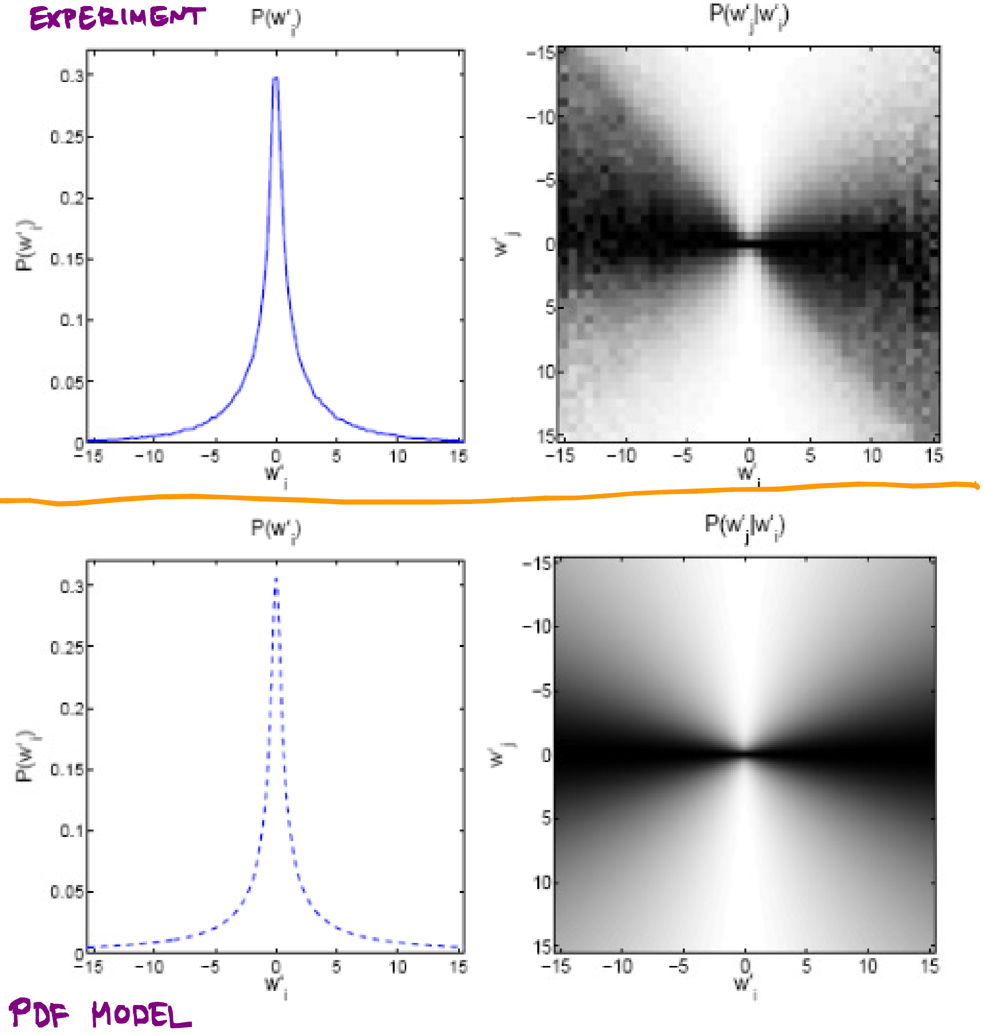
Image model in the wavelet domain

[Malo & Laparra 10]

$$P_{w'}(w') = \frac{1}{(2\pi)^{\frac{d}{2}} |\Sigma(w')|^{\frac{1}{2}}} \, e^{-\frac{1}{2} w'^{T} \cdot \Sigma(w')^{-1} \cdot w'}$$

where:

$$\Sigma_{ii}(w') = \left(\beta_i^{\gamma} + \sum_j h_{ij} |w_j|^{\gamma}\right)^{\frac{2}{\gamma}}$$

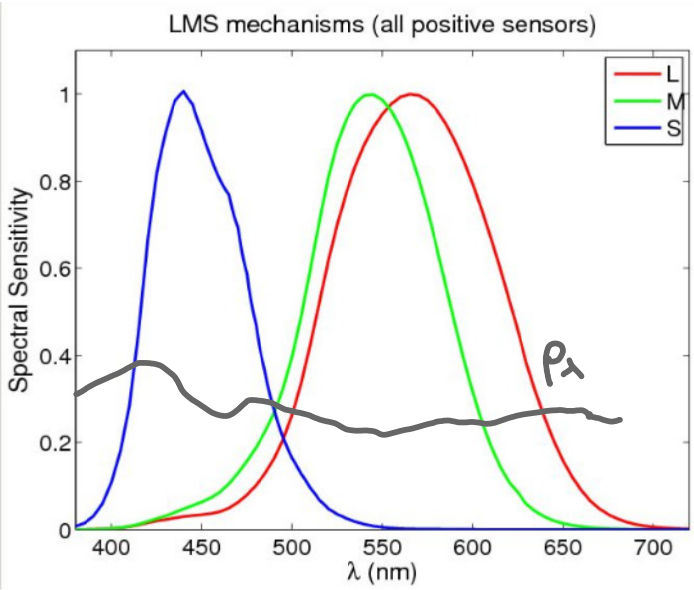


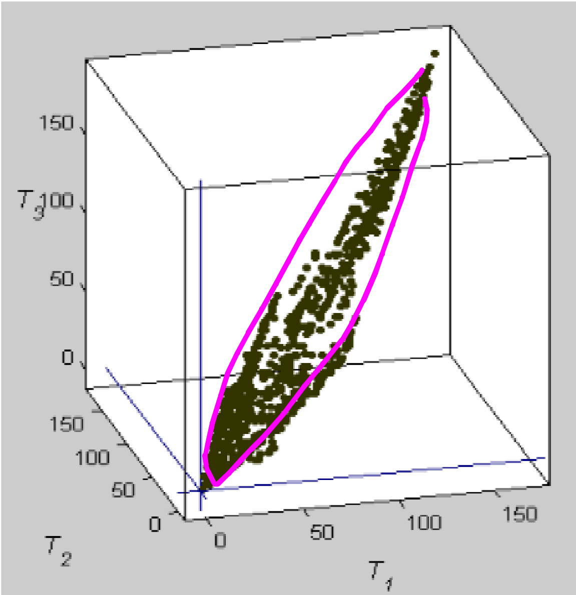
② BASIC STATISTICAL FEATURES OF IMAGES AND COLORS

2.4 Properties of natural colors

Colors are smooth!

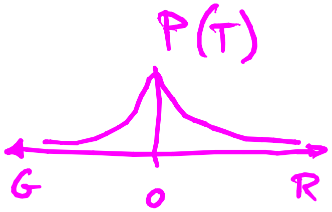
Physically feasible spectral sensitivities + Physical restrictions on reflectance

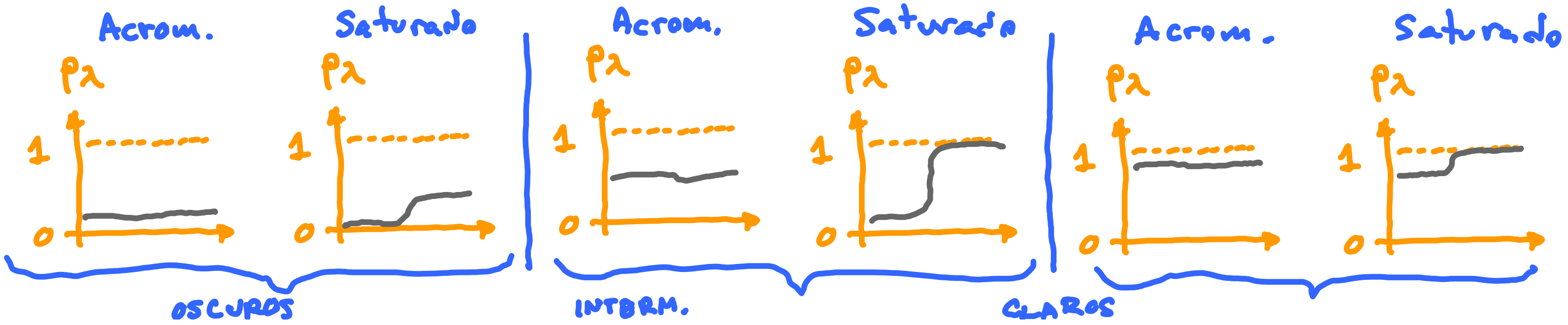




⇓

PARTICULAR COLOR MANIFOLD

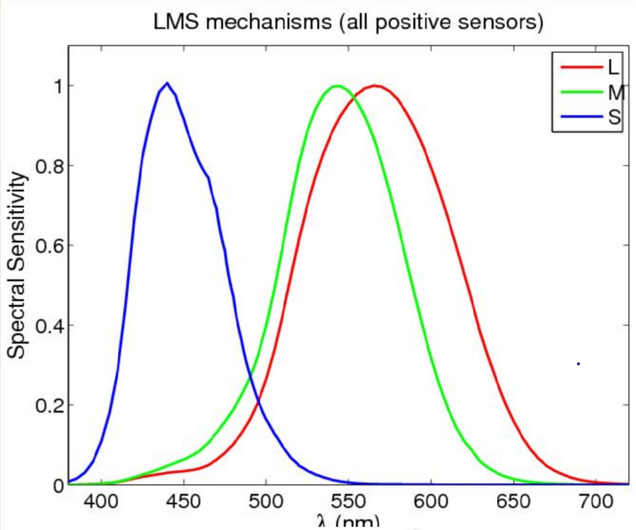


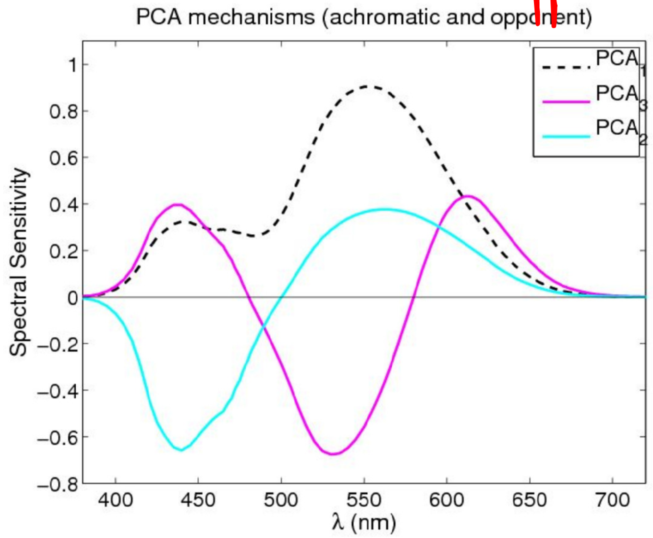


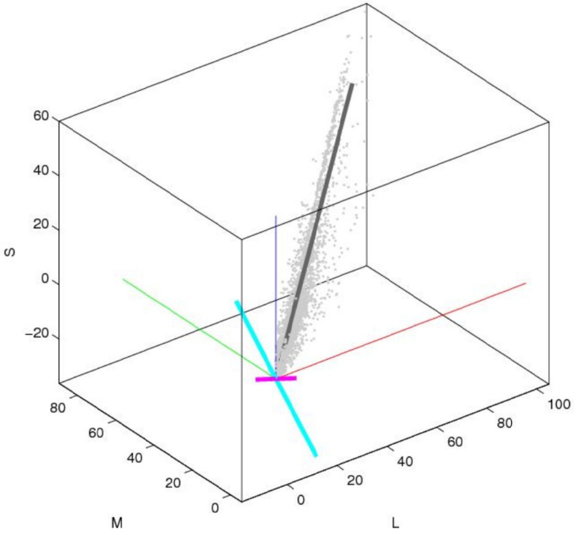
② BASIC STATISTICAL FEATURES OF IMAGES AND COLORS

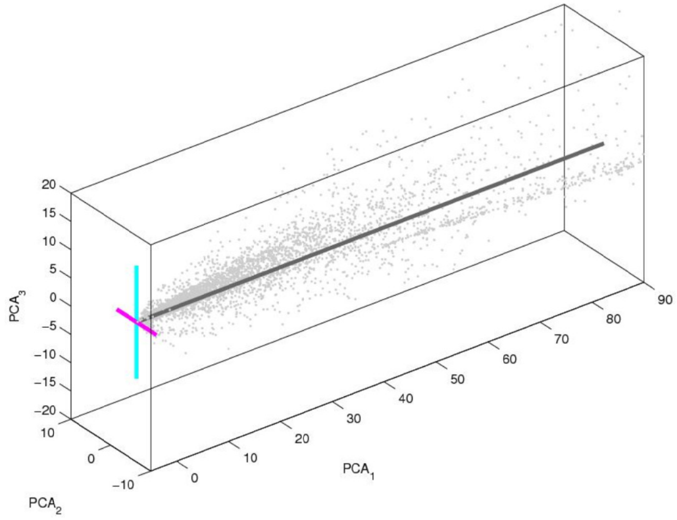
2.4 Properties of natural colors

PCA gives rise to
- Achromatic
- Opponent channels


LMS mechanisms (all positive sensors)


PCA mechanisms (achromatic and opponent)

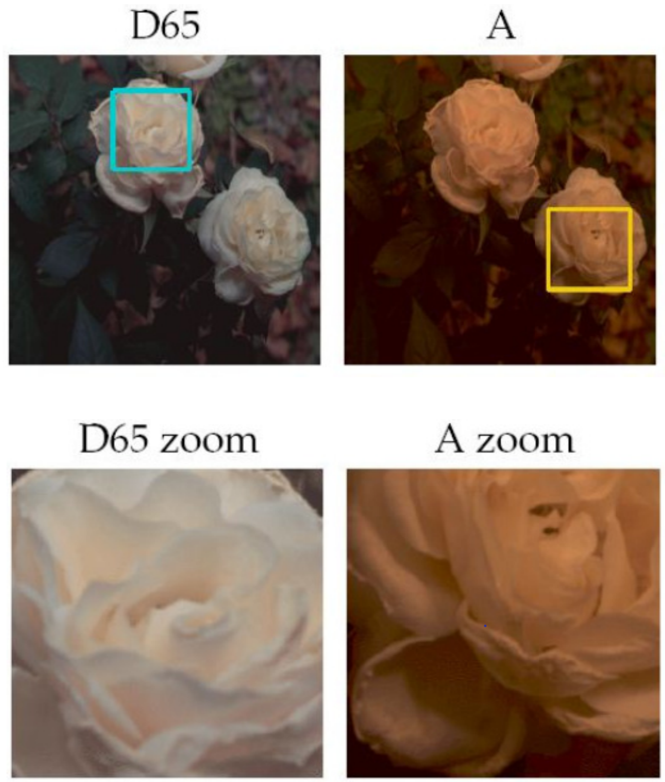


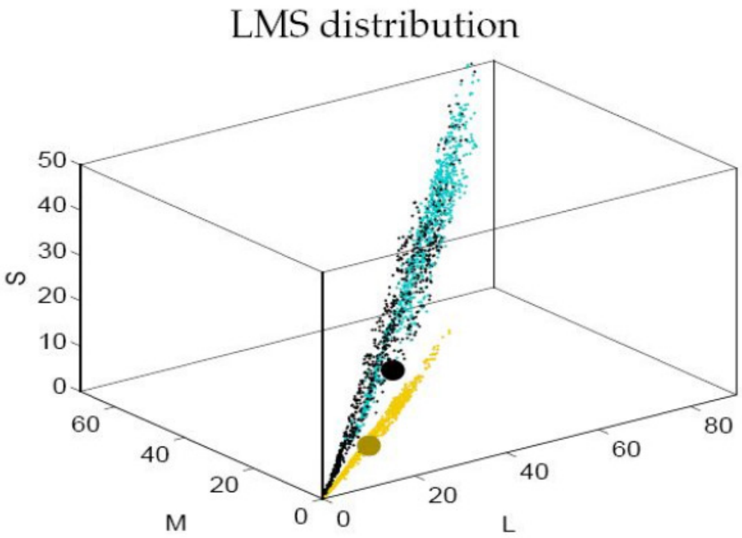


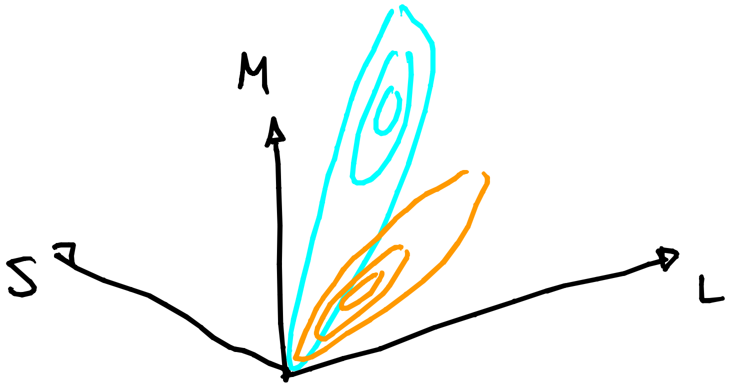
② BASIC STATISTICAL FEATURES OF IMAGES AND COLORS

2.4 Properties of natural colors

Colors (almost) linearly change with illumination





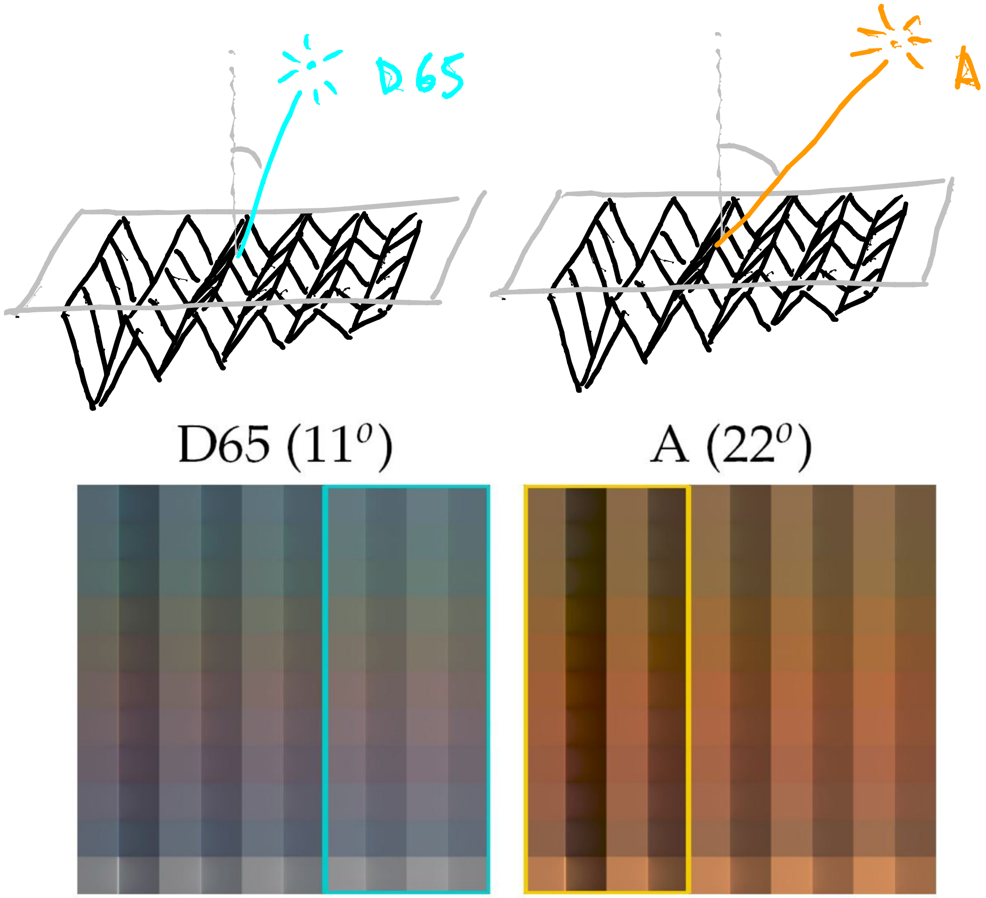


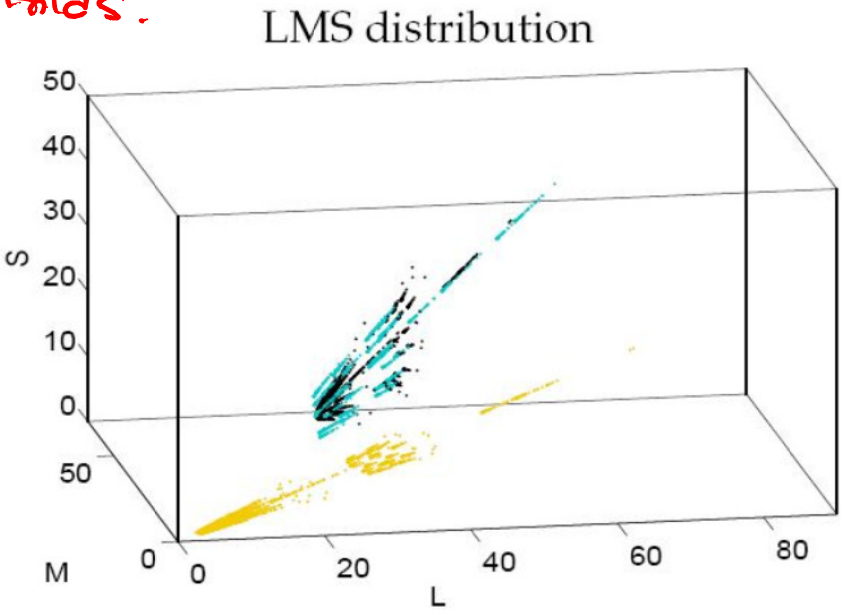
② BASIC STATISTICAL FEATURES OF IMAGES AND COLORS

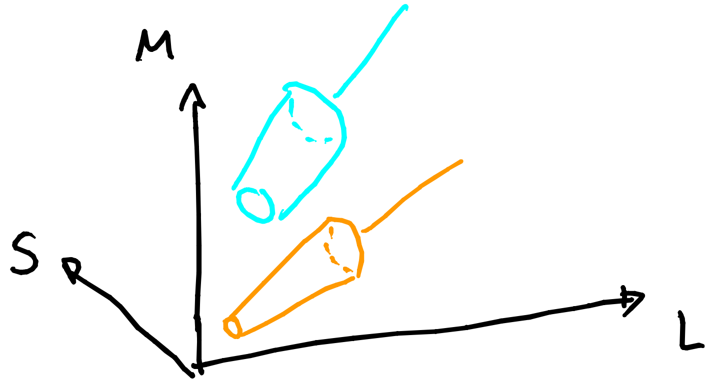
2.4 Properties of natural colors

HOWEVER, colors are more than that.

Non-linear changes in the manifolds.




LMS distribution



② BASIC STATISTICAL FEATURES OF IMAGES AND COLORS CONCLUSION

IMAGES

- * Images are smooth ⇒
  - · Low dimensional
  - · Particular PCA basis functions and spectrum (PCA ~ DCT and spectrum $\propto \frac{1}{f}$)
- * Images are non-Gaussian and have distinct features such as edges ⇒ linear ICA ~ Wavelets
- * HOWEVER, images are more than that: there are relations after linear ICA. ⇒ NON LINEARITIES REQUIRED

COLORS

- * Color components (from trichromatic sensors) are correlated in particular ways
  - ⇒ Opponent channels are better than trichromatic channels
  - Particular marginal PDFs
- * Changes in illumination are approximately linear
- * Geometry issues ⇒ NON-LINEARITIES ARE REQUIRED

QUESTIONS:

- * How does your brain process texture and color information?
- * What unsupervised learning tells us about the optimal way to do it?
- * Are the above related?
- * Can we use this stuff for image processing?

SUMMARY : MANIFOLD LEARNING IN VISION

① Something about me: research agenda THE BARLOW HYPOTHESIS

② Statistical features of images and colors

③ **Phenomenology of human vision (in textures and color)**

④ Two unsupervised learning algorithms SPCA and RBIG

⑤ Neuroscience from statistics and viceversa

⑥ Applications in image processing

- Coding
- Denoising
- Classification
- Synthesis
- Color constancy
- Image quality

③ BASIC PHENOMENOLOGY OF HUMAN VISION

3.1 Texture vision mechanisms

- Linear behavior: frequency selectivity and linear receptive fields
- Non-linear behavior: contrast masking and non-linear response in V1
- The standard (phenomenological) model: wavelets + div. Norm.

3.2 Color vision mechanisms

- Linear behavior: {
  - LMS sensors and ATD opponent channels
  - Linear adaptation transforms
- Non-linear behavior: non-linearities of ATD channels
- The standard (phenomenological) model: ATD + non-linearities

③ BASIC PHENOMENOLOGY OF HUMAN VISION

3.1 Texture vision mechanisms LINEAR behavior: frequency selectivity

f_{noise} ~ 3 cpd

f_{noise} ~ 6 cpd

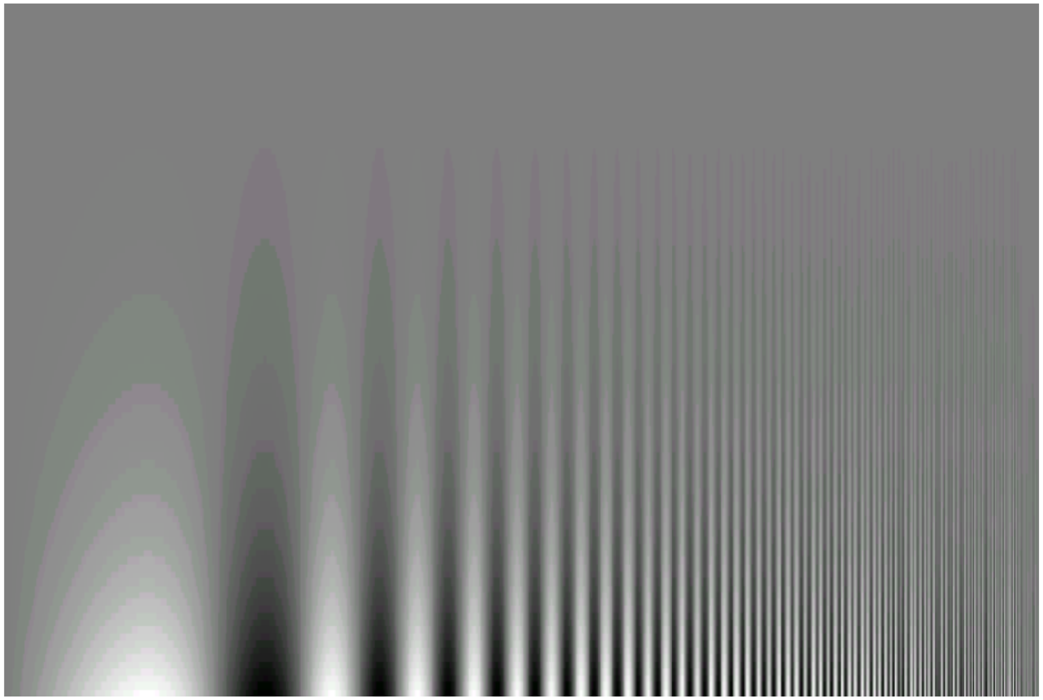
f_{noise} ~ 12 cpd

f_{noise} ~ 24 cpd

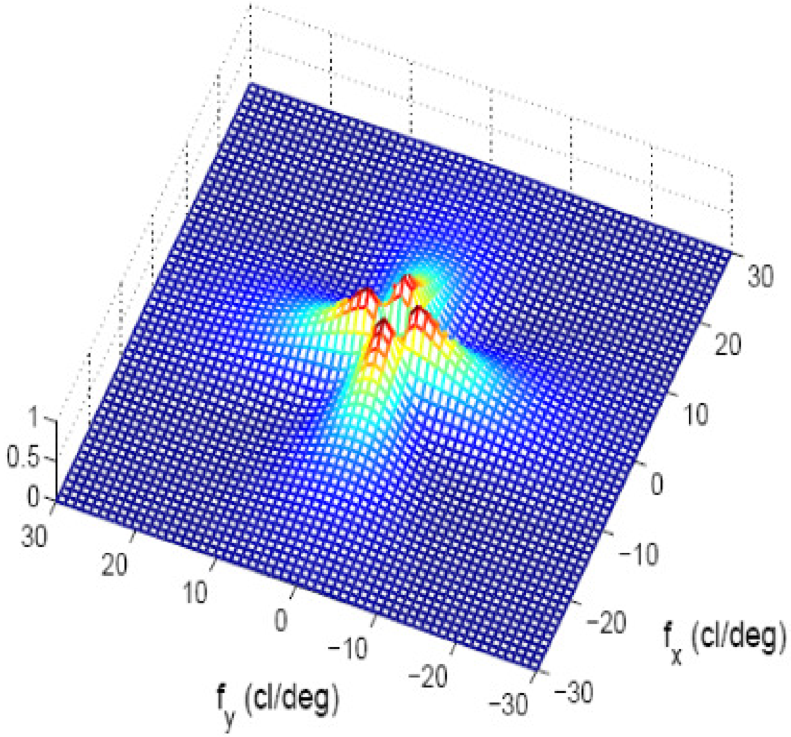
③ BASIC PHENOMENOLOGY OF HUMAN VISION

3.1 Texture vision mechanisms

LINEAR behavior: frequency selectivity



Robson test



(Linear) Standard Spatial Observer

[Watson & Ramirez 2000]

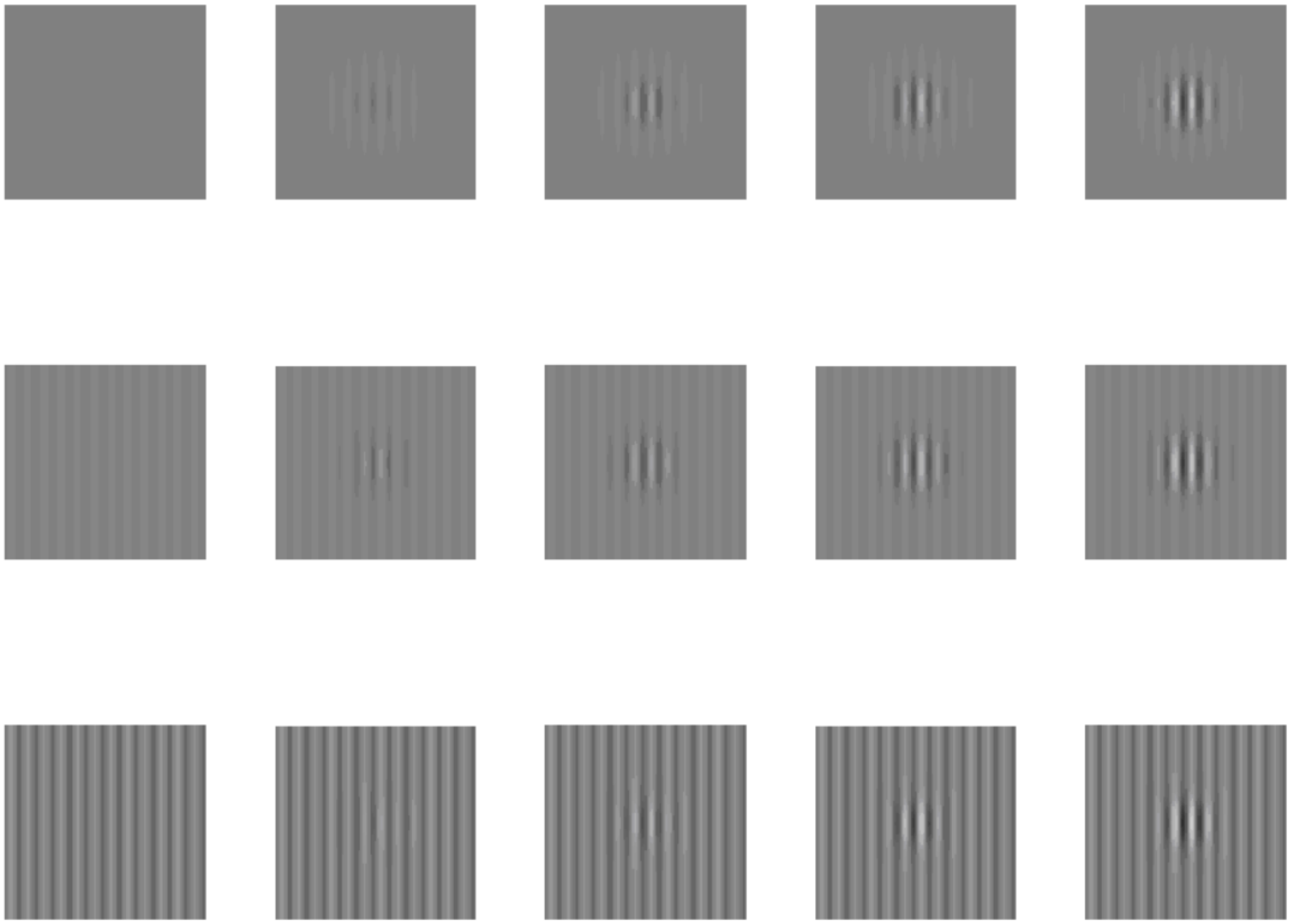
[Mullen 85]

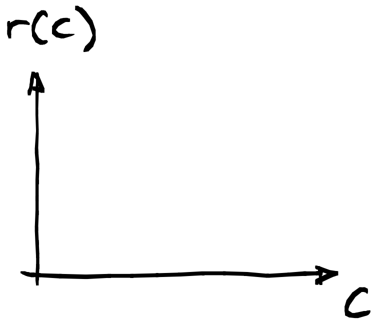
③ BASIC PHENOMENOLOGY OF HUMAN VISION

3.1 Texture vision mechanisms NON-LINEAR behavior: contrast masking

[Watson 97]

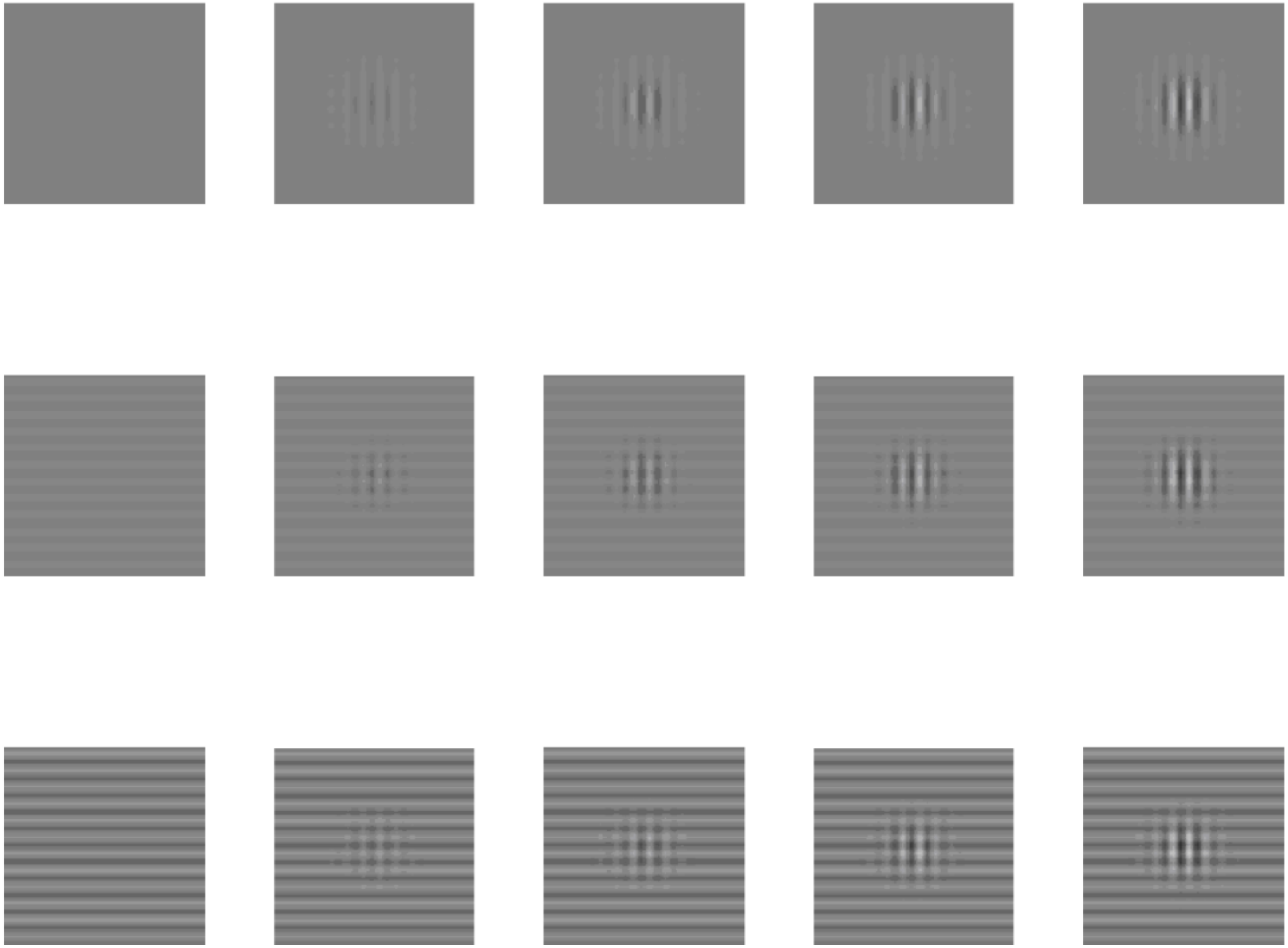
[Heeger 94]

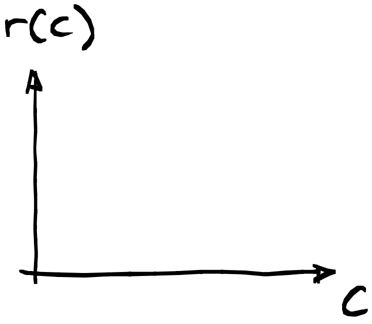




③ BASIC PHENOMENOLOGY OF HUMAN VISION

3.1 Texture vision mechanisms NON-LINEAR behavior: contrast masking





③ BASIC PHENOMENOLOGY OF HUMAN VISION

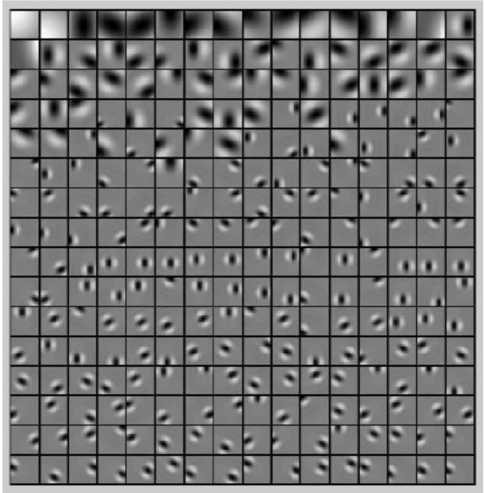
[Watson 97]

[Laparra et al. 10]

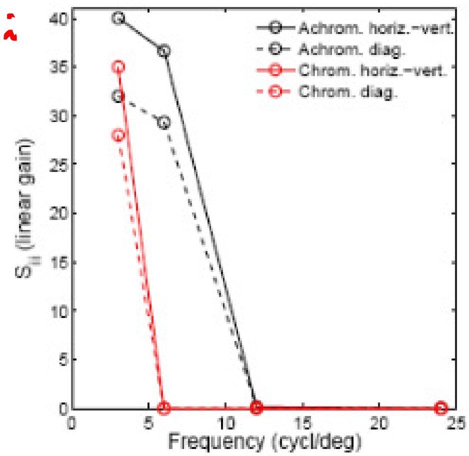
[Malo & Laparra 10]

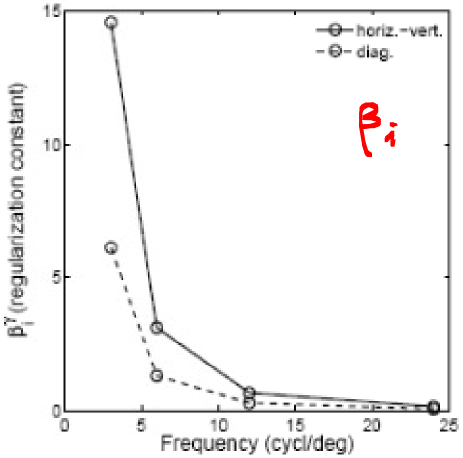
3.1 Texture vision mechanisms: PHENOMENOLOGICAL MODEL

V1 Model: $x \xrightarrow{T} w \xrightarrow{S} w' \xrightarrow{R} r$



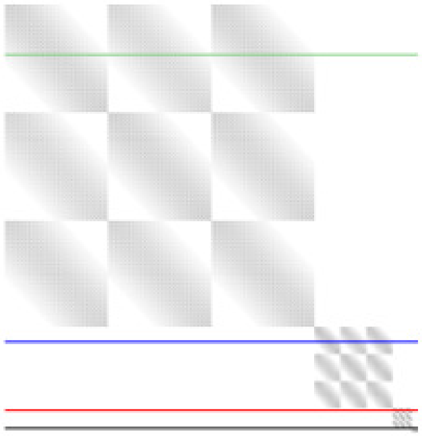
S_{ii}

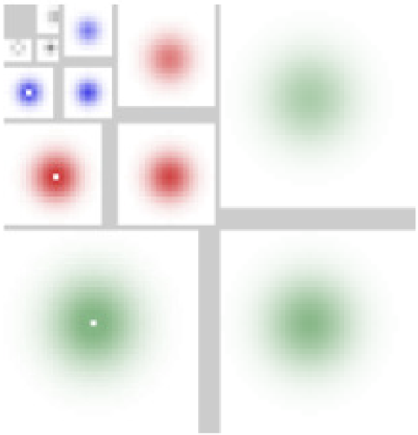




$$r_i = \text{sign}(w_i') \cdot \frac{|w_i'|^\gamma}{\beta_i^\gamma + \sum_{j=1}^{d} H_{ij} |w_j'|^\gamma}$$

H_{ij}





γ Excitation / inhibition exponent = 1'7

β_i Saturation constants

H_{ij} Interaction Kernel

③ BASIC PHENOMENOLOGY OF HUMAN VISION

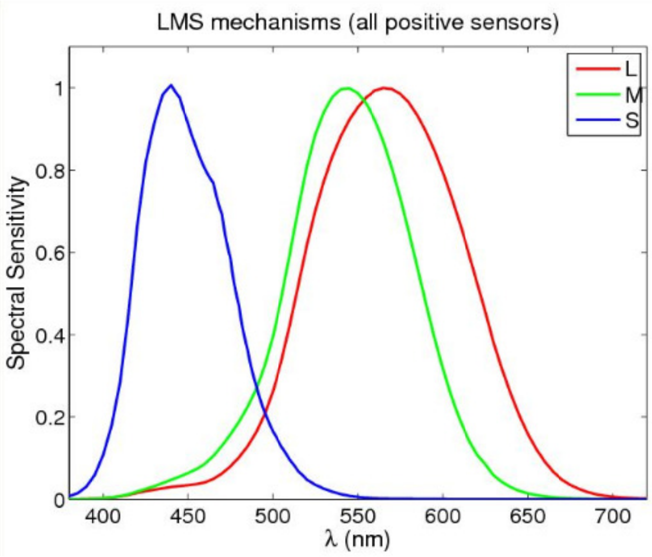
3.2 Color vision mechanisms: LINEAR BEHAVIOR

- LMS sensors
- ATD sensors

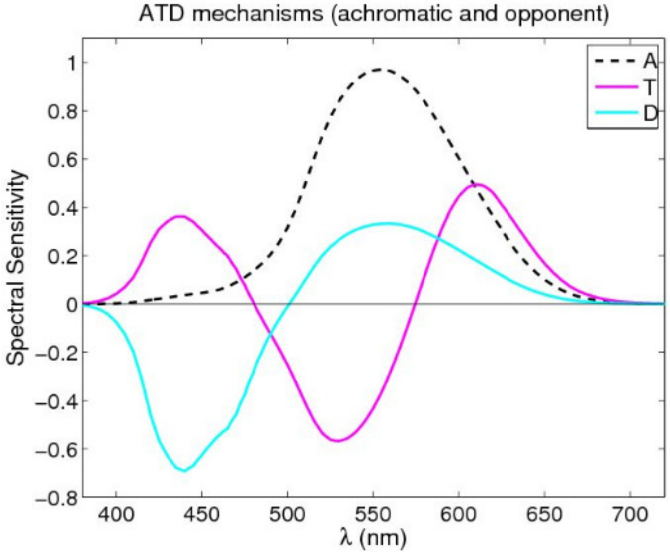
$$T_i(c) = \int c(\lambda)\, \bar{T}_i(\lambda)\, d\lambda$$

[Stockman & Brainard 09]

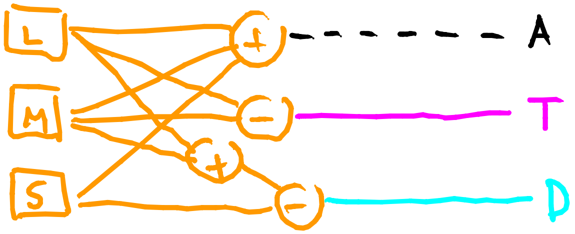
[Malo et al. COLORLAB 02]



LMS mechanisms (all positive sensors)



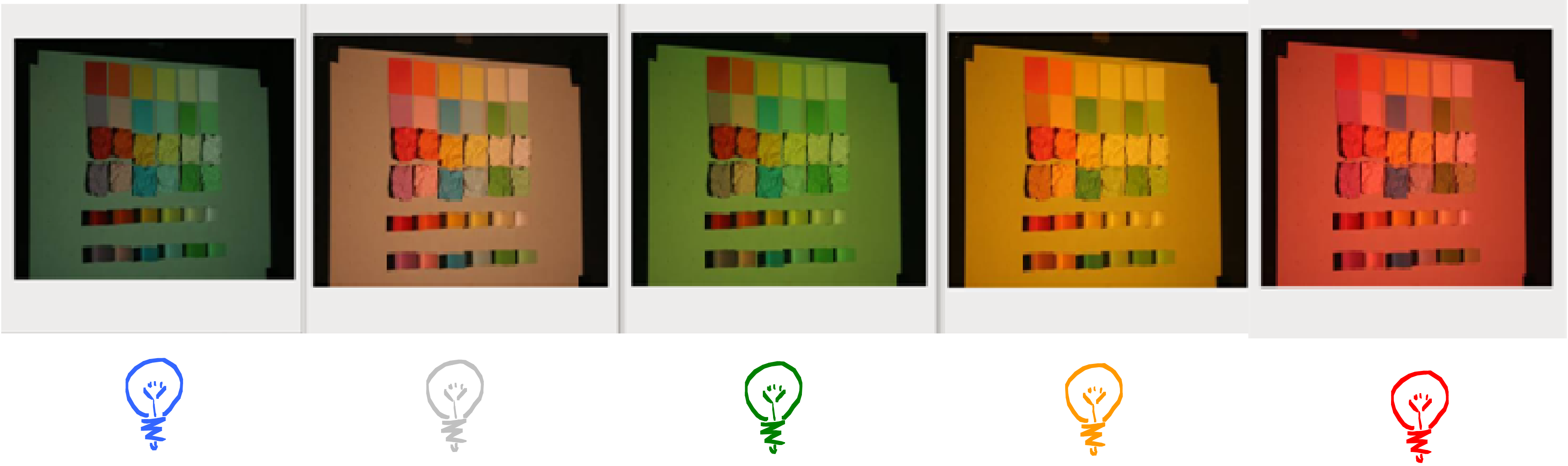
ATD mechanisms (achromatic and opponent)



$$\begin{pmatrix} A \\ T \\ D \end{pmatrix} \simeq \begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & 0 \\ 1 & 1 & -1 \end{pmatrix} \begin{pmatrix} L \\ M \\ S \end{pmatrix}$$

③ BASIC PHENOMENOLOGY OF HUMAN VISION

3.2 Color vision mechanisms: LINEAR BEHAVIOR linear color adaptation

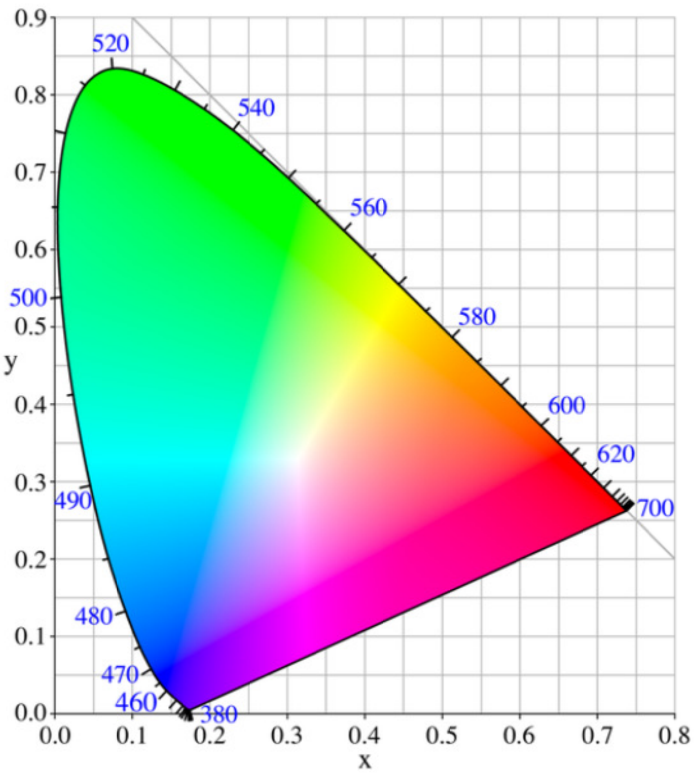


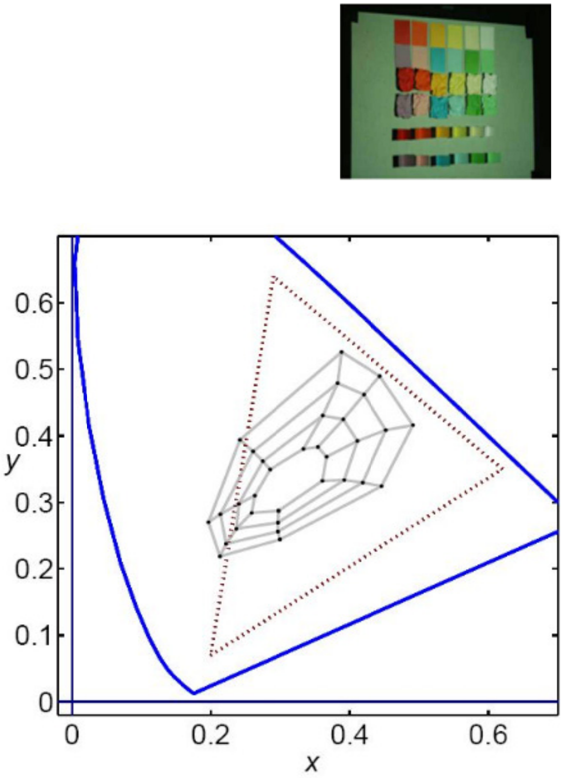
Von Kries adaptation [Brainard 05]

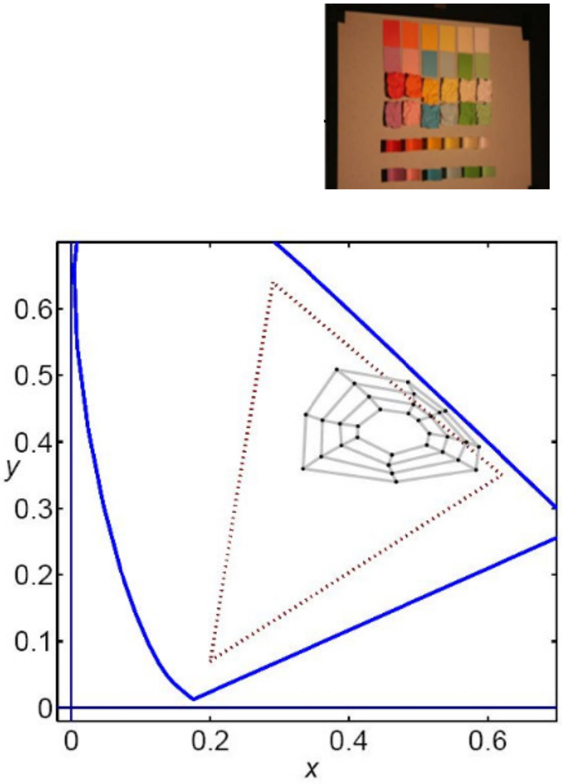
$$\begin{pmatrix} A' \\ T' \\ D' \end{pmatrix} = \begin{pmatrix} \frac{1}{A(w)} & 0 & 0 \\ 0 & \frac{1}{T(w)} & 0 \\ 0 & 0 & \frac{1}{D(w)} \end{pmatrix} \begin{pmatrix} A \\ T \\ D \end{pmatrix}$$

③ BASIC PHENOMENOLOGY OF HUMAN VISION

3.2 Color vision mechanisms: LINEAR BEHAVIOR linear color adaptation







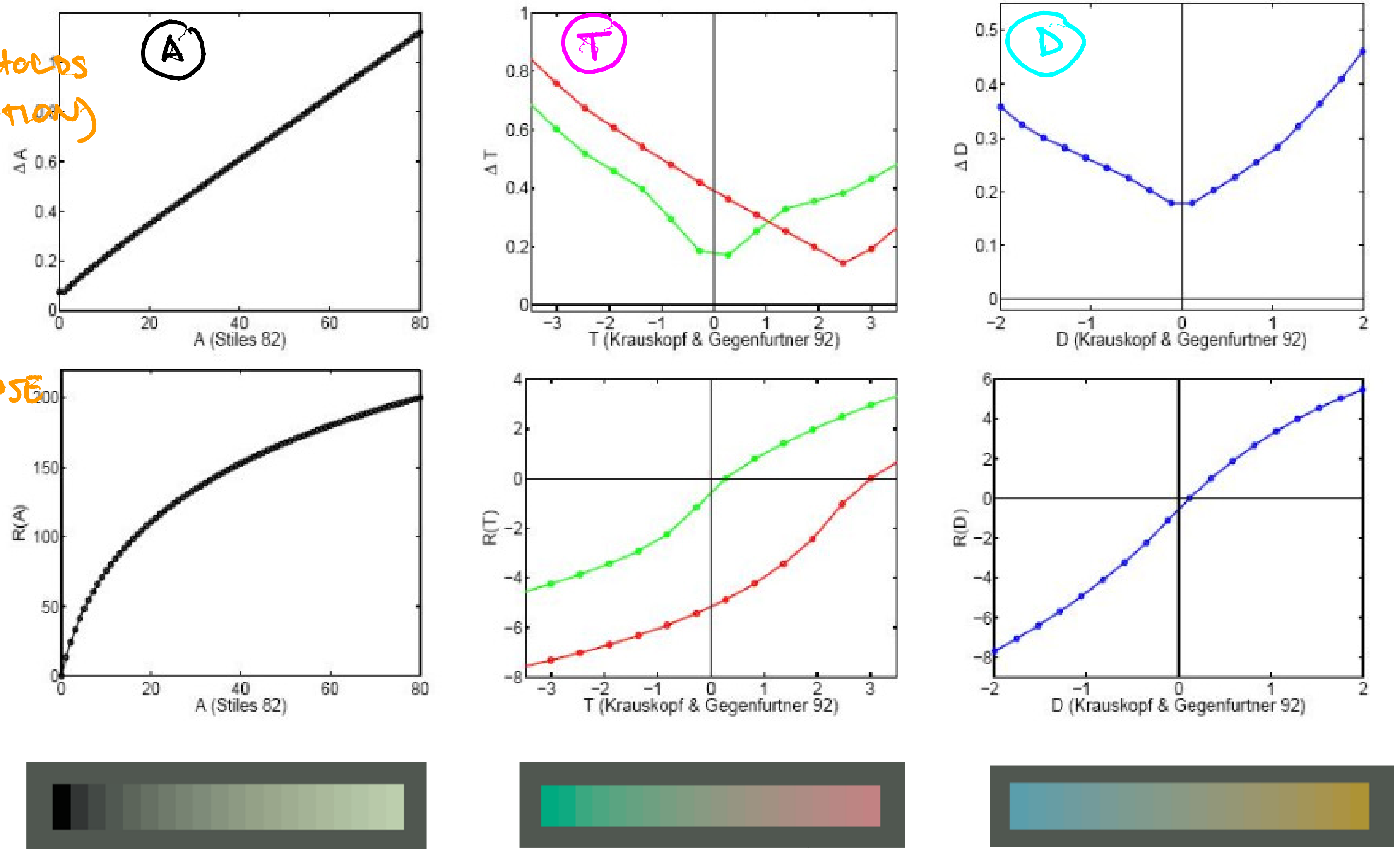
[Data from Luo et al. 91]

③ BASIC PHENOMENOLOGY OF HUMAN VISION

3.2 Color vision mechanisms: NON-LINEAR BEHAVIOR

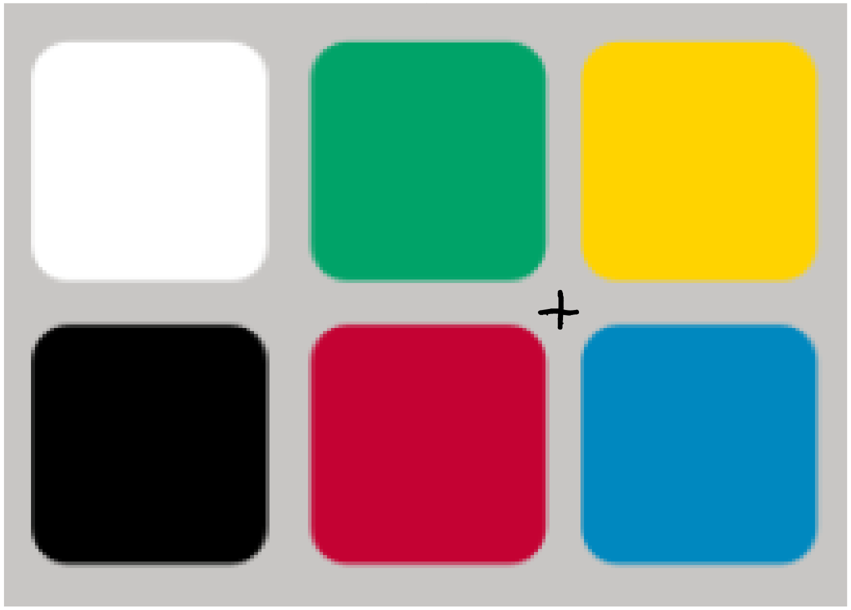
THRESHOLDS (RESOLUTION)

RESPONSE



③ BASIC PHENOMENOLOGY OF HUMAN VISION

3.2 Color vision mechanisms: Summary: Post effects.



③ BASIC PHENOMENOLOGY OF HUMAN VISION

3.2 Color vision mechanisms: PHENOMENOLOGICAL MODEL [Malo et al. COLORLAB 02]

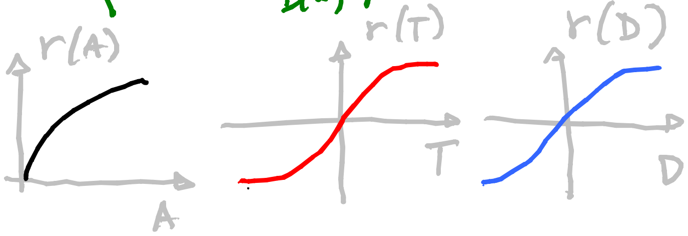
$$\begin{tikzcd}
C(\lambda) \arrow[r, "\tau"] & T_{LMS} \arrow[r, "M"] & T_{ATD} \arrow[r, "V"{name=V}] & T'_{ATD} \arrow[r, "R"] & r_{ATD} \\
W(\lambda) \arrow[to=V]
\end{tikzcd}$$

$\tau \equiv$ Trichromatic broadband linear sensors $T_i(C) \approx \int C(\lambda)\,\bar{T}_i(\lambda)\,d\lambda$

$M \equiv$ Opponent space $\left\{ \begin{array}{l} \text{Achromatic} \\ \text{Red-green} \\ \text{yellow-blue} \end{array} \right.$ $\quad M \sim \begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & 0 \\ 1 & 1 & -1 \end{pmatrix}$

$V \equiv$ Von Kries linear adaptation transform $\quad V \sim \begin{pmatrix} \frac{1}{A(W)} & 0 & 0 \\ 0 & \frac{1}{T(W)} & 0 \\ 0 & 0 & \frac{1}{D(W)} \end{pmatrix}$

$R \equiv$ Non-linear saturating response



③ BASIC PHENOMENOLOGY OF HUMAN VISION: CONCLUSIONS (YOU SAW!)

TEXTURE vision mechanisms

- Linear behavior: frequency selectivity and linear receptive fields
- Non-linear behavior: contrast masking and non-linear response in V1
- The standard (phenomenological) model: wavelets + div. Norm.
  [Malo & Laparra 10]

COLOR vision mechanisms

- Linear behavior: { - LMS sensors and ATD opponent channels
  - Linear adaptation transforms }
- Non-linear behavior: non-linearities of ATD channels
- The standard (phenomenological) model: ATD + non-linearities
  [Malo et al. COLORLAB 02]

QUESTIONS: How does this relates to image & color statistics as revealed by unsuperv. Learn.?
Is this useful in image processing?

SUMMARY : MANIFOLD LEARNING IN VISION

1. Something about me : research agenda THE BARLOW HYPOTHESIS
2. Statistical features of images and colors
3. Phenomenology of human vision (in textures and color)
4. **Two unsupervised learning algorithms SPCA and RBIG**
5. Neuroscience from statistics and viceversa
6. Applications in image processing
   - Coding
   - Denoising
   - Classification
   - Synthesis
   - Color constancy
   - Image quality

④ TWO UNSUPERVISED LEARNING ALGORITHMS

4.1 Non-linear ICA ≡ infomax solutions are not unique

- Meaningful
- Meaningless

4.2 Sequential Principal Curves Analysis (SPCA)

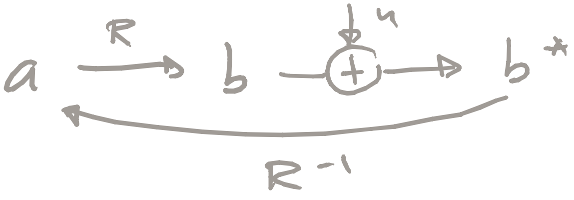
- Basic fact: Unfolding and equalization lead to component independence
- The algorithm: unfold and equalize through a sequence of PCs
- Practical behavior

4.3 Rotation-Based Iterative Gaussianization (RBIG)

- Estimation of PDFs
- Projection pursuit and our approach
- Practical behavior

④ TWO UNSUPERVISED LEARNING ALGORITHMS

4.1 Non-linear ICA ≡ infomax solutions are not unique



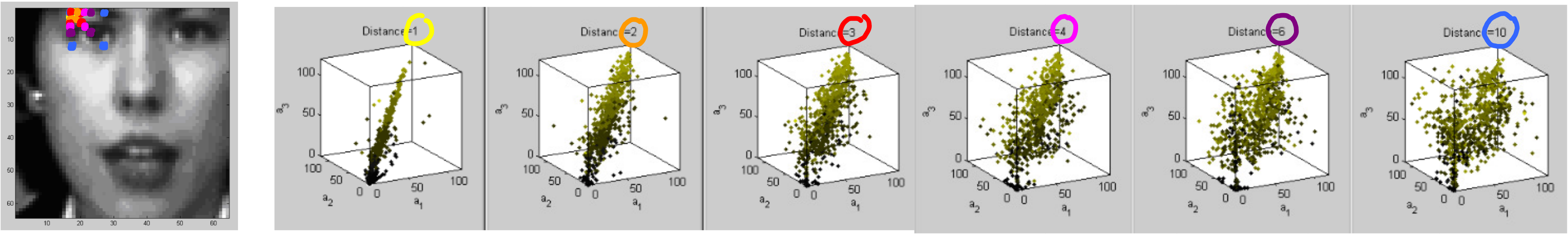
[Bell & Sejnowski 95] ⇒ $|\nabla R| \propto p(a)$

[Twer & MacLeod 01] ⇒ $|\nabla R| \propto p(a)^{\frac{1}{3}}$

} ⇒ Restrict. on $|\nabla R|$ BUT. No unique prescription for R

[Hyvarinen 03]

Example: Random scrambling of pixels ⇒ ~ Gaussianization



QUESTION: Do we want that? (independence at any cost?)

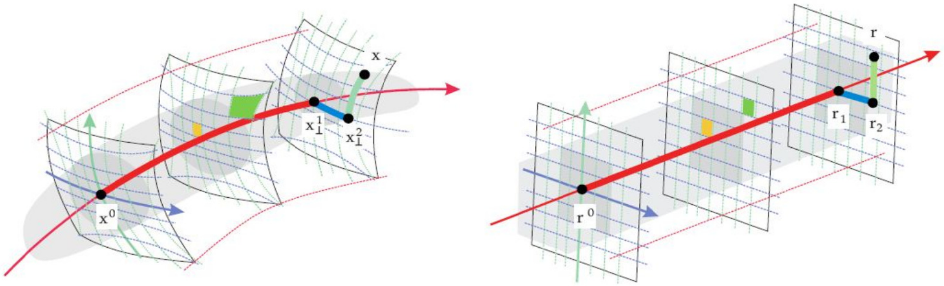
④ TWO UNSUPERVISED LEARNING ALGORITHMS

4.1 Non-linear ICA ≡ infomax solutions are not unique

We tried 2 approaches.

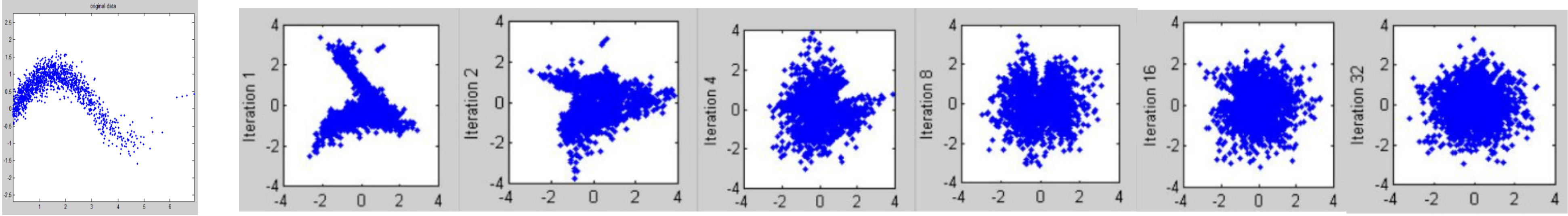
* Meaningful: Sequential Principal Curves Analysis (SPCA)

[Laparra et al. 12?
Malo & Gutierrez 06]



* Meaningless: Rotation-Based Iterative Gaussianization (RBIG)

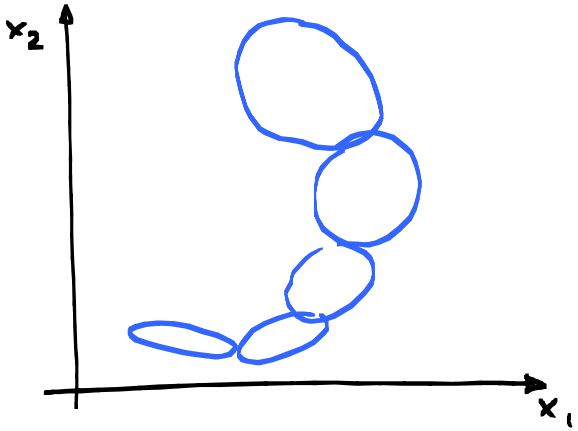
[Laparra et al. 11]

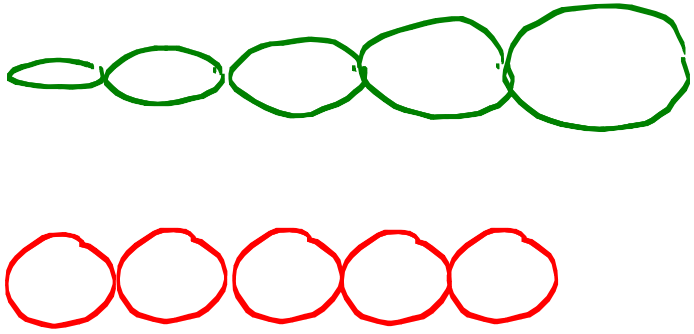


④ TWO UNSUPERVISED LEARNING ALGORITHMS

4.2 Sequential Principal curves analysis (SPCA): UNFOLDING AND EQUALIZAT.

$$\begin{tikzcd} x \arrow[r, shift left, "R"] & r \arrow[l, shift left, "R^{-1}"] \end{tikzcd}$$





* Basic asumption: mixture of clusters

- Roweis, Saul, Hinton NIPS 02
- Verbeek et al. ICANN 02
- Teh & Roweis NIPS 03
- Brand NIPS 03
- * Malo & Gutierrez Network 06

$$r = \int_x \nabla R(x)\,dx \quad , \quad \nabla R(x) = A(x)$$

* Proposed approach

  - * UNFOLDING
  - * LOCAL METRIC

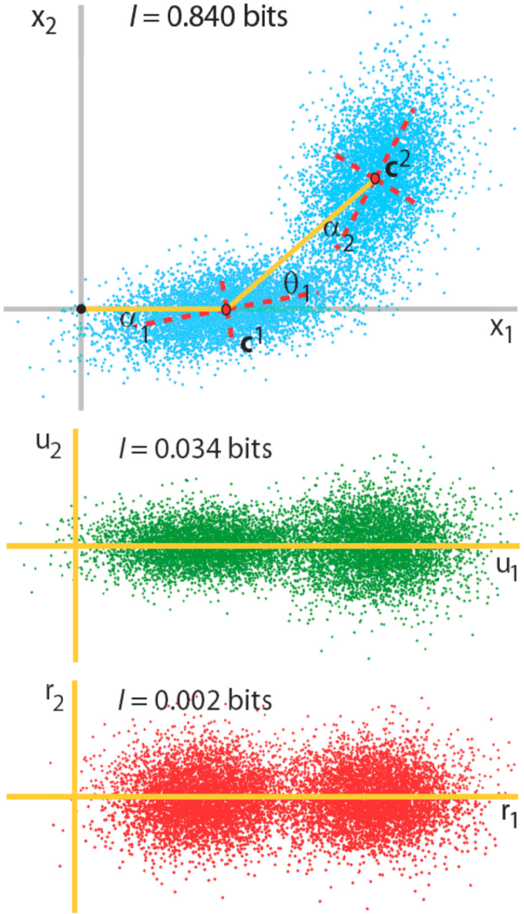
$$\nabla R(x) = \underbrace{\Delta(x)}_{\text{LOCAL METRIC}} \cdot \underbrace{T(x)}_{\text{ROTATIONS ALONG PRINCIPAL CURVES}} = p(x)^{\gamma} \cdot T(x)$$

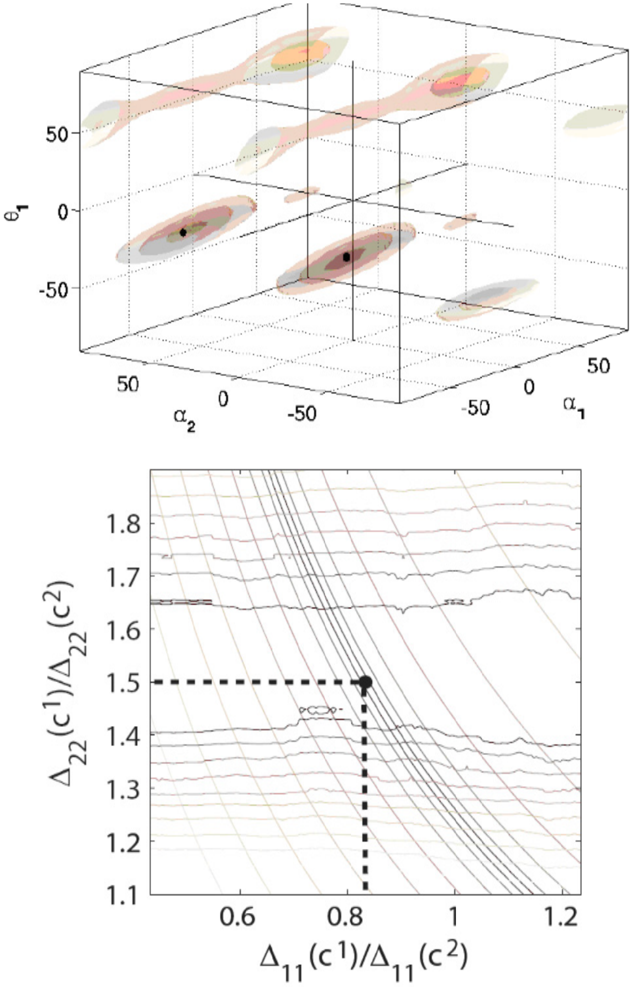
$$\boxed{r(x) = \int_{PC} \Delta(x) \cdot T(x)\,dx}$$

④ TWO UNSUPERVISED LEARNING ALGORITHMS

4.2 Sequential Principal curves analysis (SPCA):

Unfolding & Equalization do reduce multiinformation (the goal!)

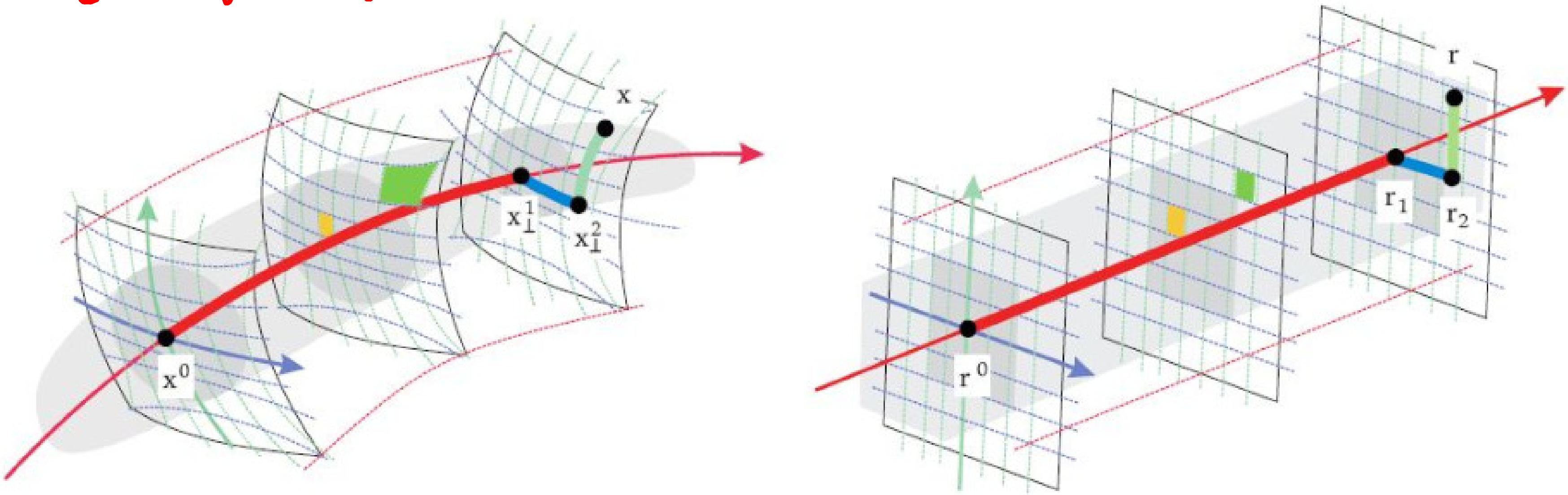




④ TWO UNSUPERVISED LEARNING ALGORITHMS

4.2 Sequential Principal curves analysis (SPCA)

d-dimensional invertible representation with independent coefficients by integrating $\Delta(x)T(x)$ on a SEQUENCE of PCs.



NOTE:
You can use whatever PC algorithm e.g.
[Hastie 00, Delicado 01]

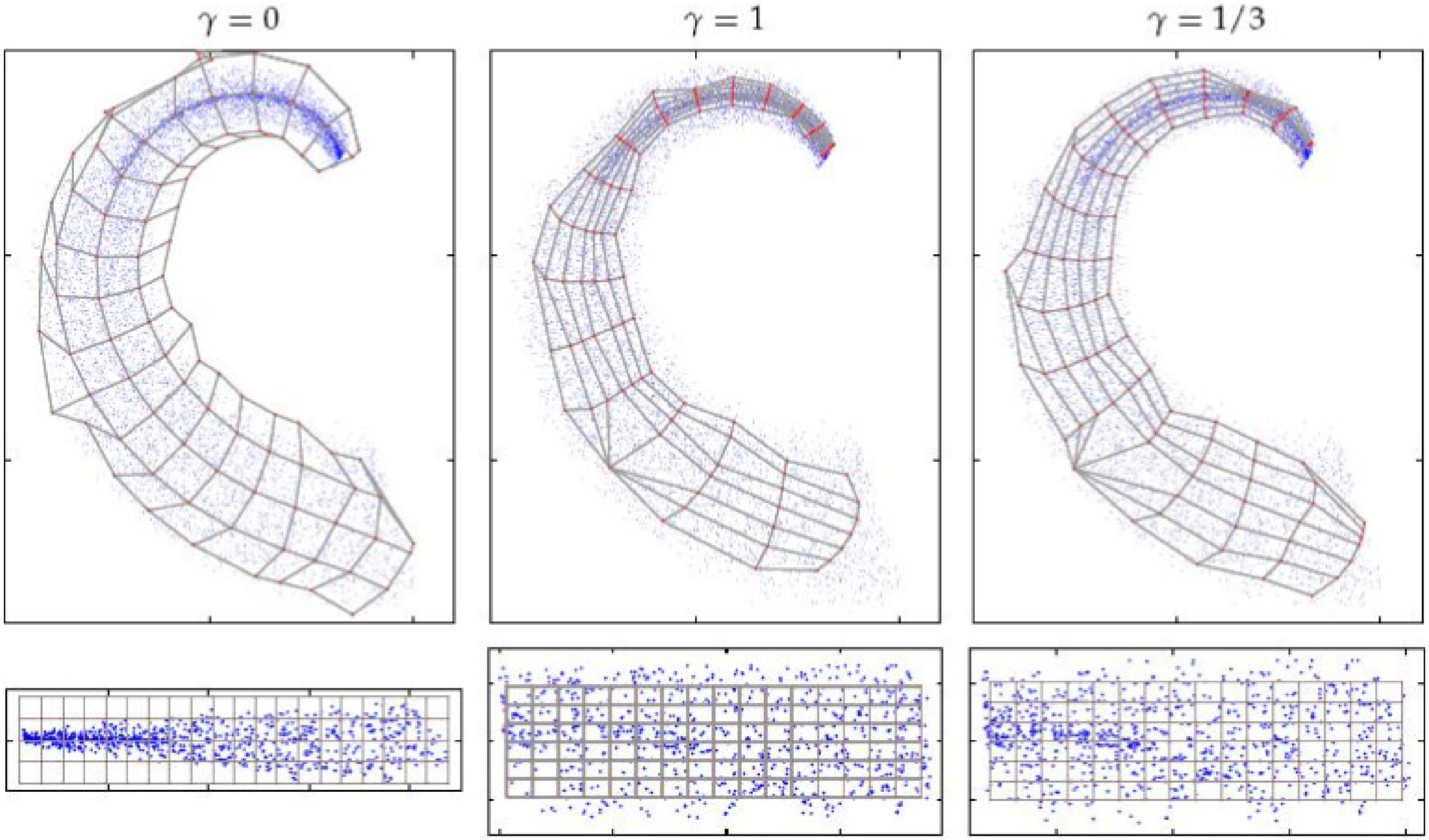
$$\begin{tikzcd} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \arrow[r, bend left, "R"] & \begin{pmatrix} r_1 \\ r_2 \\ r_3 \end{pmatrix} \arrow[l, bend left, "R^{-1}"] \end{tikzcd}$$

$$r(x) = \int_{x_0}^{x} p(x)^{\gamma} T(x) \, dx$$

through PCs

④ TWO UNSUPERVISED LEARNING ALGORITHMS

4.2 Sequential Principal curves analysis (SPCA): Practical behavior



Infomax

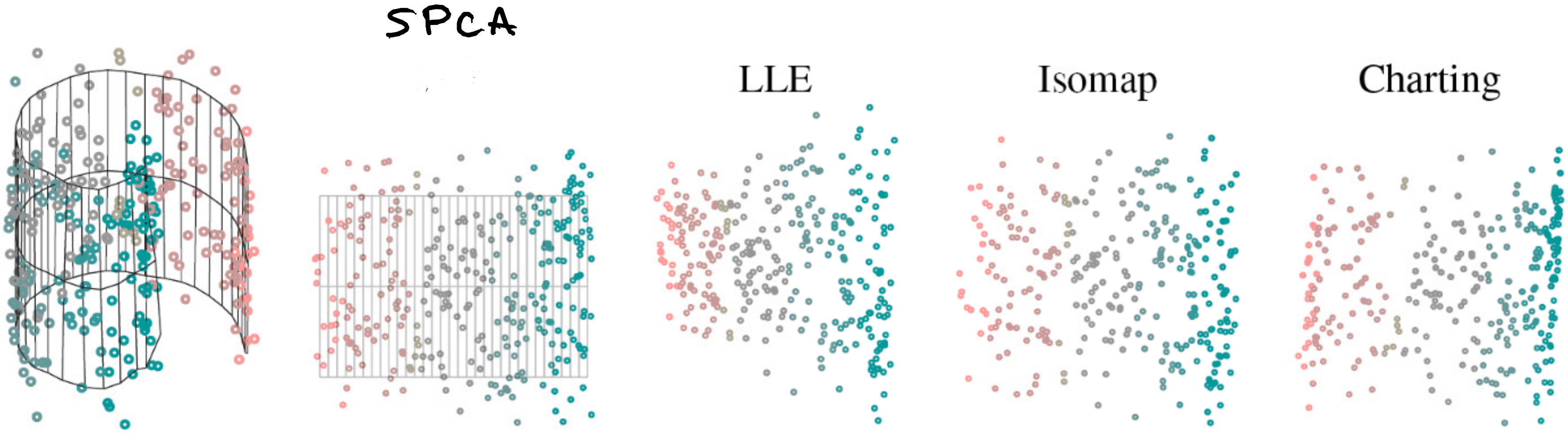
Error minimiz.

$\gamma = 0$ $\gamma = 1$ $\gamma = 1/3$

RMSE	0.55	0.56	0.53
MI (bits)	0.27	0.05	0.06

④ TWO UNSUPERVISED LEARNING ALGORITHMS

4.2 Sequential Principal curves analysis (SPCA): Practical behavior

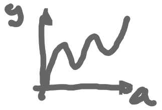


④ TWO UNSUPERVISED LEARNING ALGORITHMS

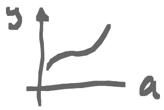
4.3 Rotation-based iterative Gaussianization (RBIG)

* The problem of PDF estimation: PDFs in domains related by F, are related

$$a \xrightarrow{F} y$$

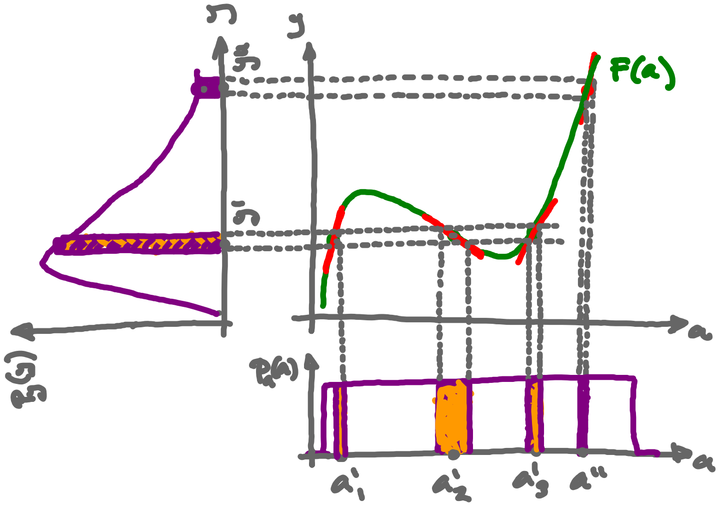


$$P_y(y) = \sum_{a \in F^{-1}(y)} P_a(a) \, |\nabla F(a)|^{-1}$$



$$P_y(y) = P_a(a) \, |\nabla F(a)|^{-1} \quad \text{con} \quad a = F^{-1}(y)$$

Intuición (en caso 1D):
La población en un punto, y', es la suma de las poblaciones en los puntos $a' = F^{-1}(y')$ pesada por la inversa de la pendiente en esos puntos $|\nabla F(a')|^{-1}$



$a'_i = F^{-1}(y')$
$a'' = F^{-1}(y'')$

* SOLUTION: look for a transform to a known PDF! (e.g. a Gaussian)

④ TWO UNSUPERVISED LEARNING ALGORITHMS

4.3 Rotation-based iterative Gaussianization (RBIG)

* Alternative routes to the Gaussian goal:

$$J = I + J_m \Rightarrow \begin{cases} \cdot \text{ Go for } I & \underline{\text{Projection Pursuit}} \quad \text{Huber 85, Chen 00} \\ \cdot \text{ Go for } J_m & \underline{\text{Our approach!}} \end{cases}$$

Projection pursuit

- Reduce I : ICA Linear transf. (R_{ICA})
- Reduce J_m : Ψ Marginal Gaussianizat.

$$\Rightarrow x^{(k+1)} = \Psi^{(k)} \odot R_{ICA}^{(k)} x^{(k)}$$

Our approach:

- Reduce J_m : Ψ
- Reduce I : ANY R !

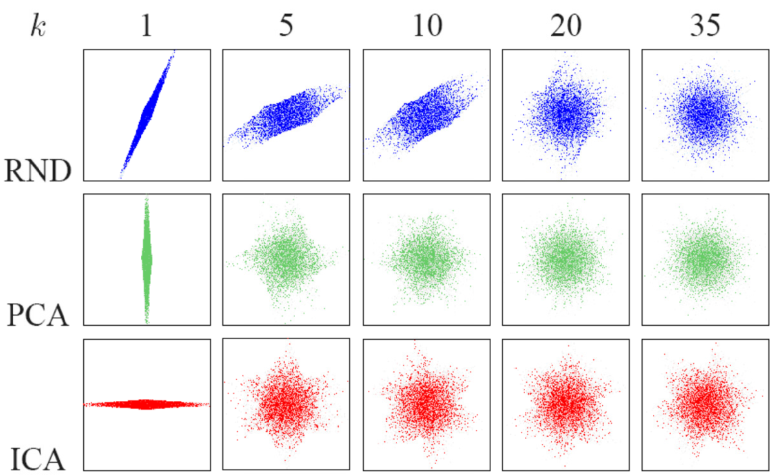
$$\Rightarrow \boxed{x^{(k+1)} = R^{(k)} \odot \Psi^{(k)} x^{(k)}}$$

$$\Delta J = J(x) - J(R\Psi(x)) = I(x) + J_m(x) - J(\Psi(x)) = \cancel{I(x)} + J_m(x) - \cancel{I(\Psi(x))} - \cancel{J_m(\Psi(x))} = J_m(x) \geq 0$$

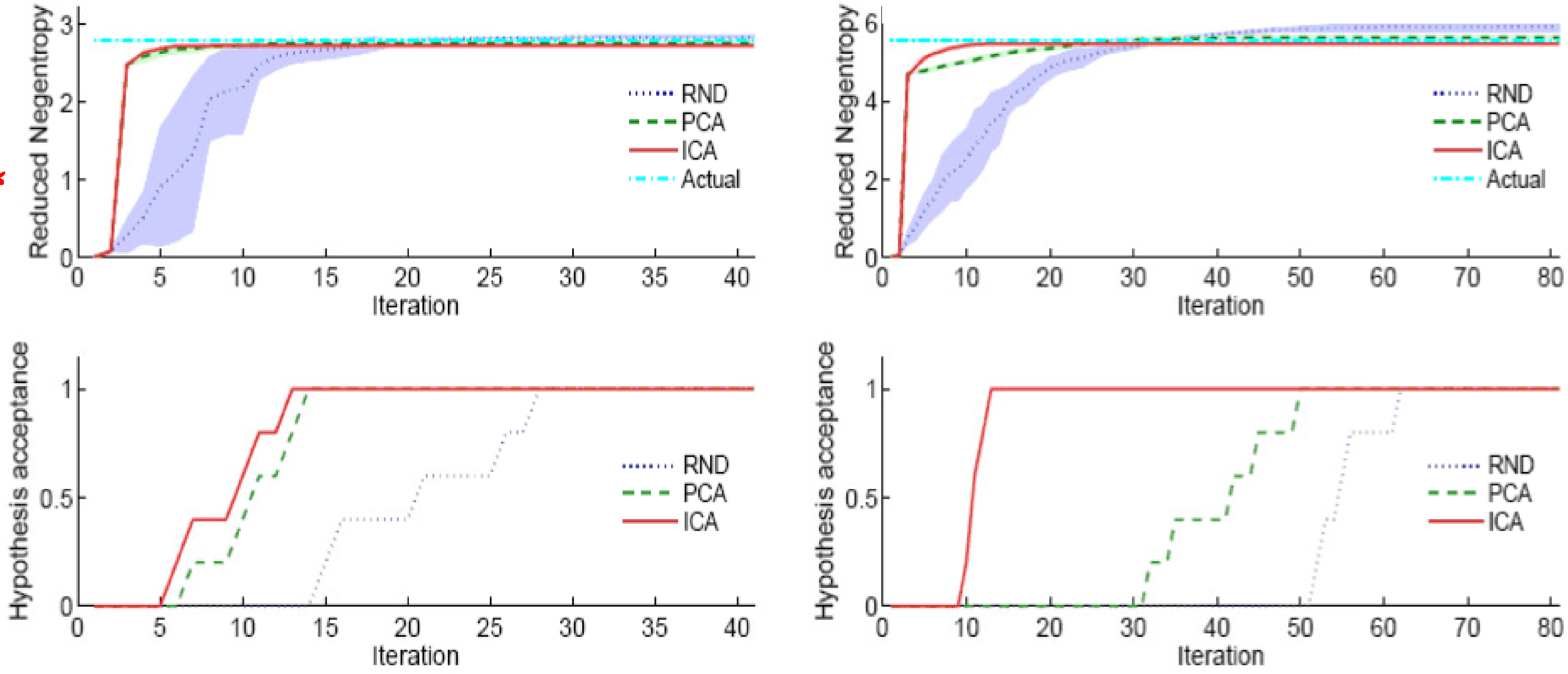
$$\Delta I = I(x) - I(R\Psi(x)) = I(\Psi(x)) - I(R\Psi(x)) = \cancel{J(\Psi(x))} - \cancel{J_m(\Psi(x))} - \cancel{J(R\Psi(x))} + J_m(R\Psi(x)) = J_m(R\Psi(x)) \geq 0$$

A.2 Projection Pursuit versus RBIG

$$\Delta J_{(k)} = J_m\left(x^{(k-1)}\right) \geq 0 \quad \forall k$$

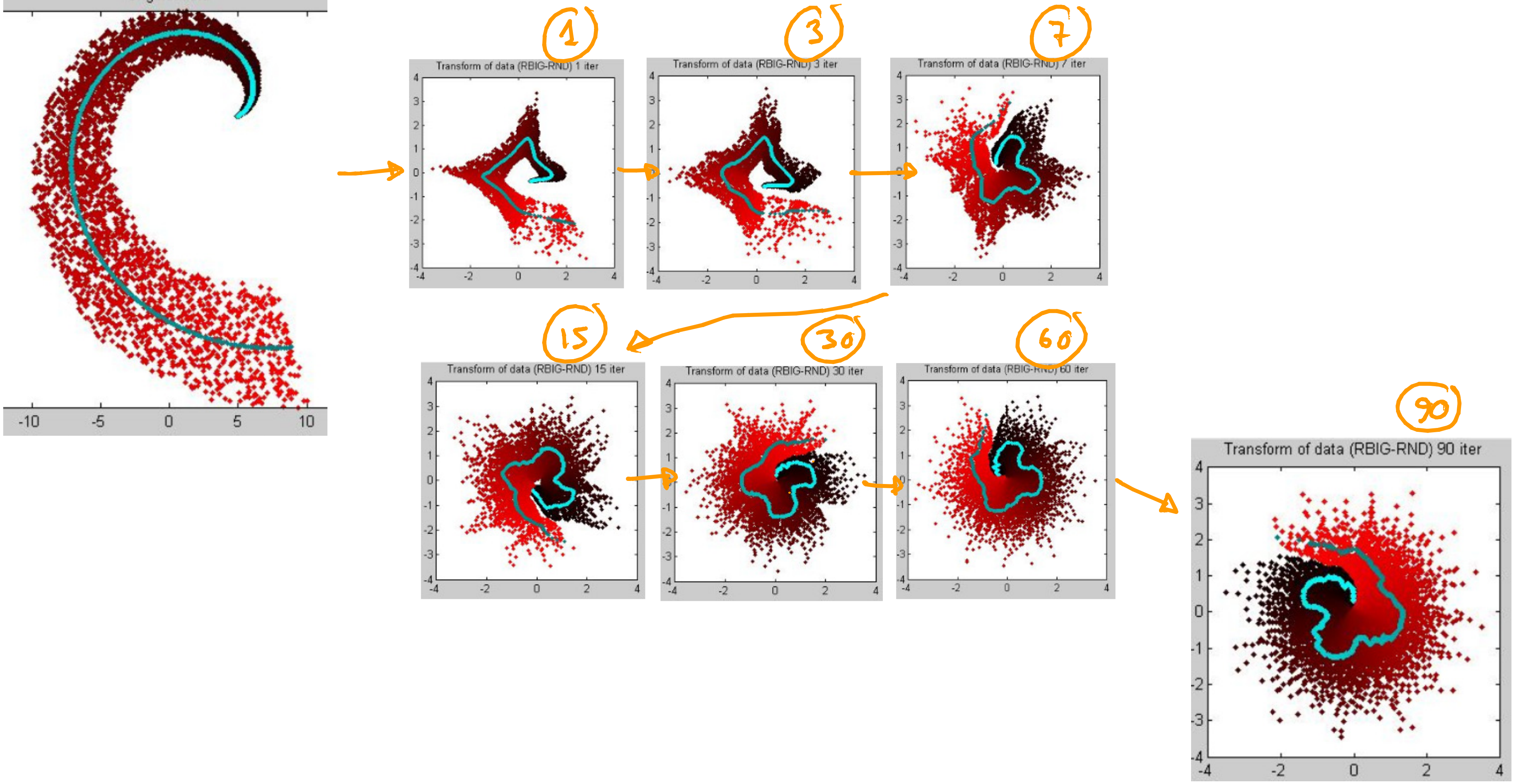


$$J = \sum_k \Delta J_{(k)}$$



④ TWO UNSUPERVISED LEARNING ALGORITHMS

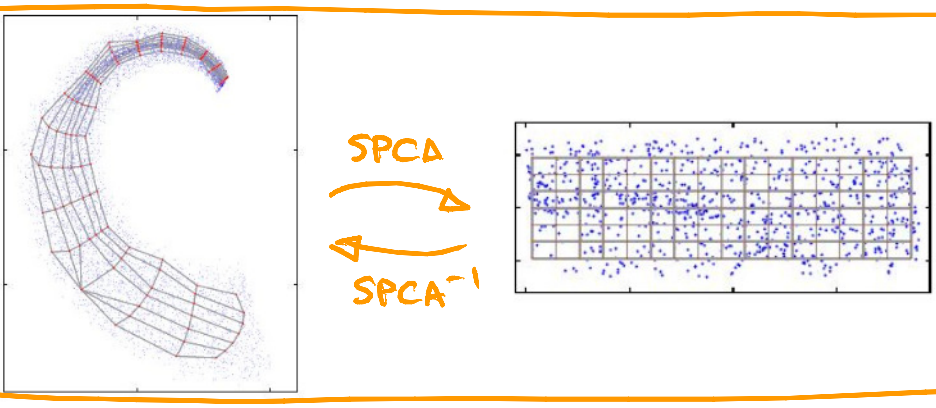
4.3 Rotation-based iterative Gaussianization (RBIG)



④ TWO UNSUPERVISED LEARNING ALGORITHMS   CONCLUSIONS

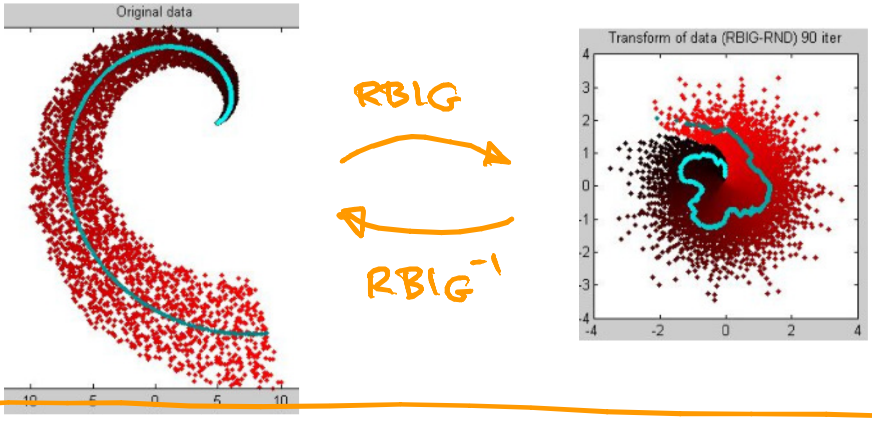
We developed two succesful NON-LINEAR ICA / PDF estimat. algorithms well suited to non-trivial image stats.

SPCA



$$r(x) = \int_{x_0}^{x} \Delta(x)\; T(x)\, dx$$
through PCs

RBIG



$$x^{(k+1)} = R^{(k)} \cdot \Psi^{(k)} \cdot x^{(k)}$$

QUESTIONS:
- What is the behavior obtained when using these on natural signals?
- Is this related to what the brain does?
- Applications in image processing?

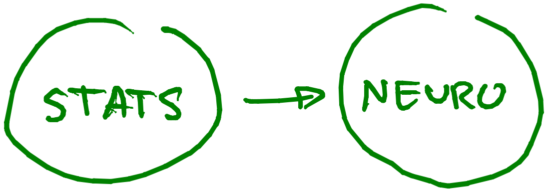
SUMMARY : MANIFOLD LEARNING IN VISION

1. Something about me: research agenda THE BARLOW HYPOTHESIS
2. Statistical features of images and colors
3. Phenomenology of human vision (in textures and color)
4. Two unsupervised learning algorithms SPCA and RBIG
5. **Neuroscience from statistics and viceversa**
6. Applications in image processing
    - Coding
    - Denoising
    - Classification
    - Synthesis
    - Color constancy
    - Image quality

⑤ NEUROSCIENCE FROM STATISTICS (AND VICEVERSA!)

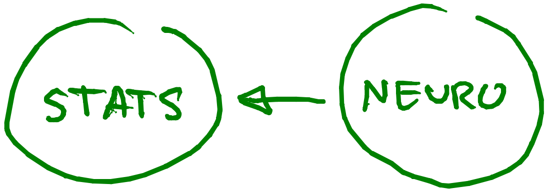
5.1 Neuroscience from statistics

- Contrast masking non-linearities by using SPCA
- Adaptation and color non-linearities by using SPCA
- Adaptation of color mechanisms using RBIG



5.2 Statistics from neuroscience (good properties of empirical models)

- The standard model of texture vision factorizes the PDF of images
- " " " " color " strongly reduces multiinform



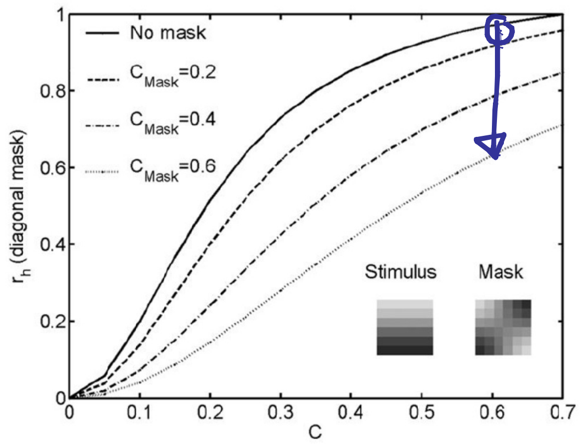
5) NEUROSCIENCE FROM STATISTICS (AND VICEVERSA!)

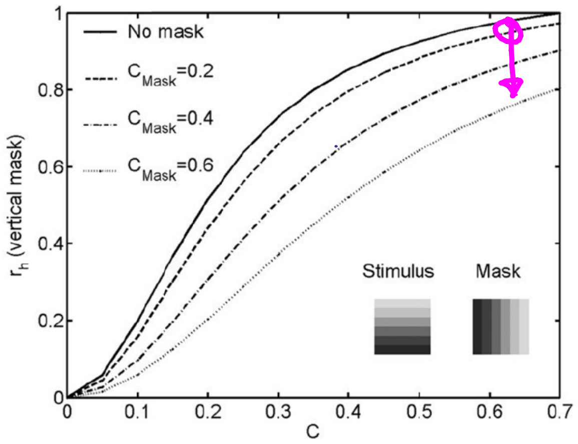
5.1 Neurosci from statistics: CONTRAST MASKING USING SPCA [Malo & Gutierr. 06]

SIMILAR MASK

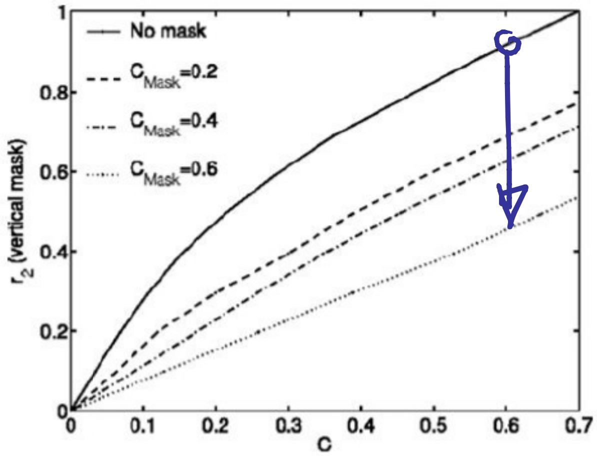
DIFFERENT MASK

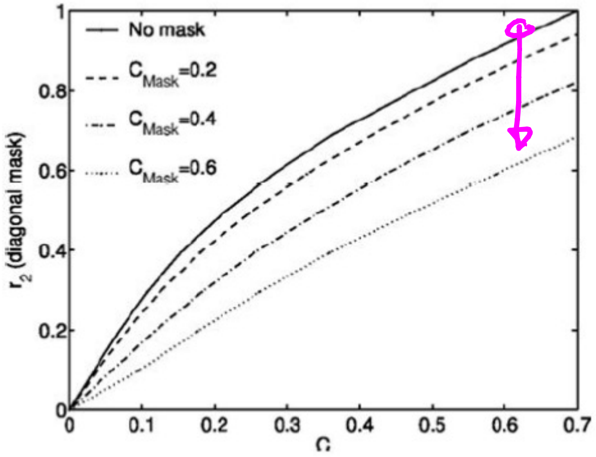
EMPIRICAL BEHAVIOR





THEORETICAL PREDICTION (SPCA)





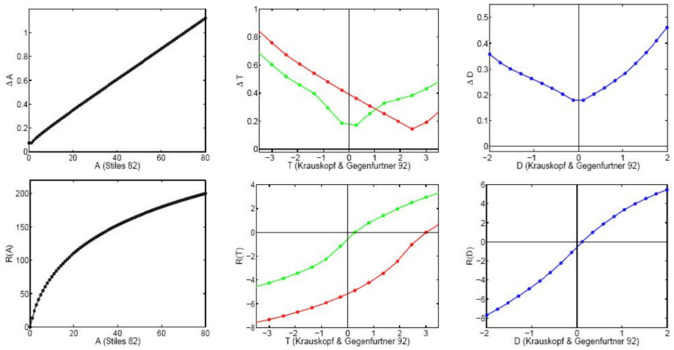
⑤ NEUROSCIENCE FROM STATISTICS (AND VICEVERSA!)

5.1 Neurosci from statistics:

COLOR NON-LINEARITIES USING SPCA

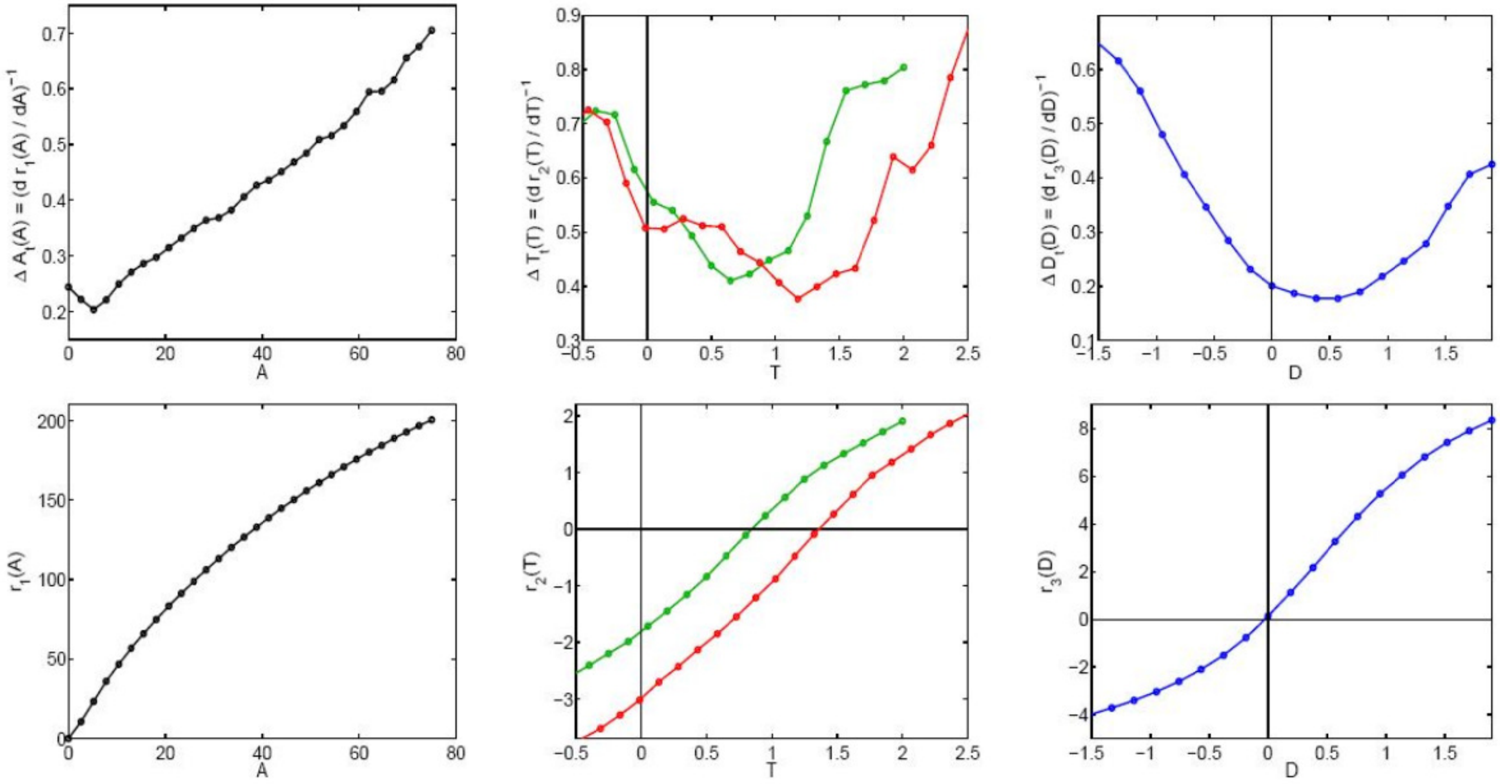
[Laparra et al. Submitted]

EMPIRIC.



THEORETIC. PREDICTION

(SPCA)



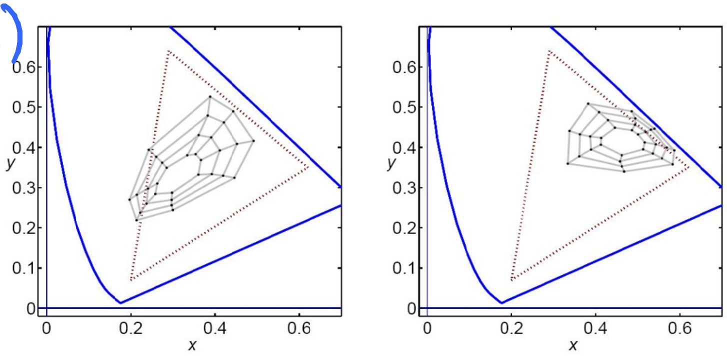
⑤ NEUROSCIENCE FROM STATISTICS (AND VICEVERSA!)

5.1 Neurosci from statistics:

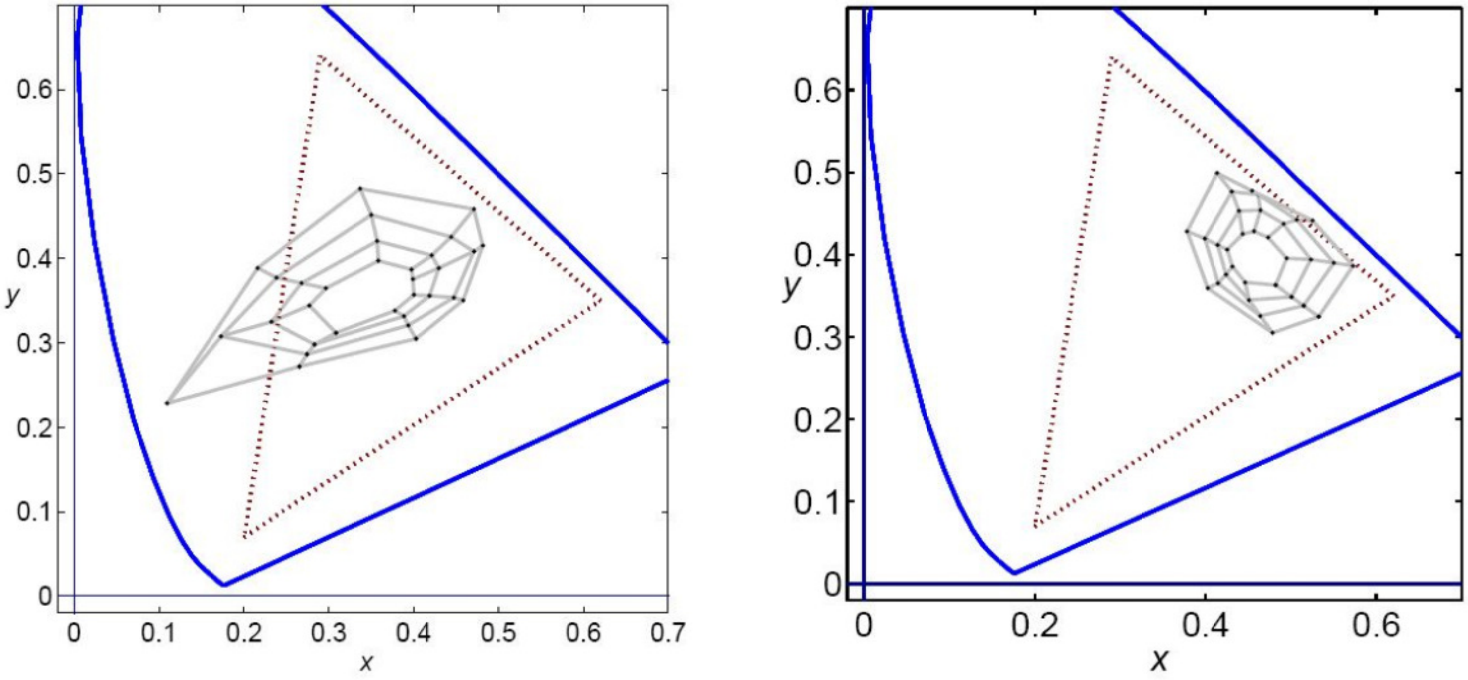
COLOR ADAPTATION USING SPCA

[Laparra et al. Submitted]

EMPIRIC.

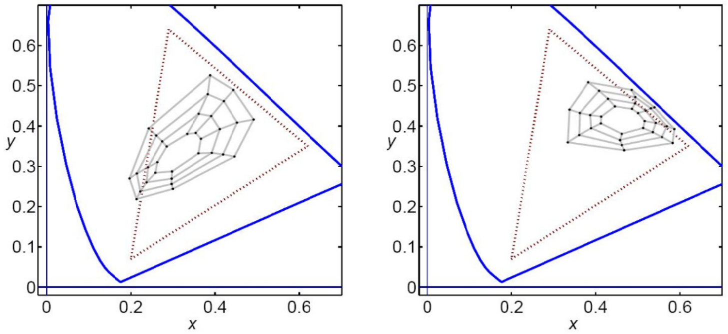


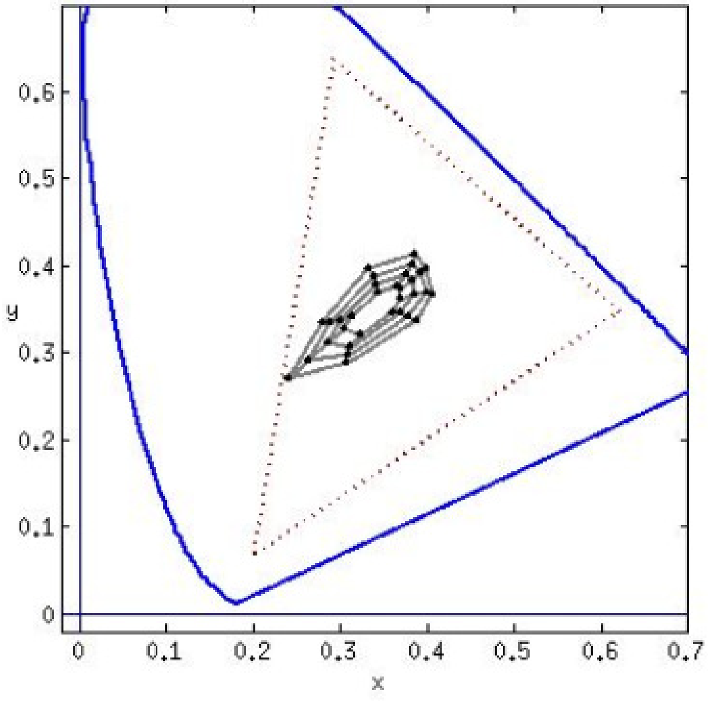
THEORETICAL PREDICTION (SPCA)

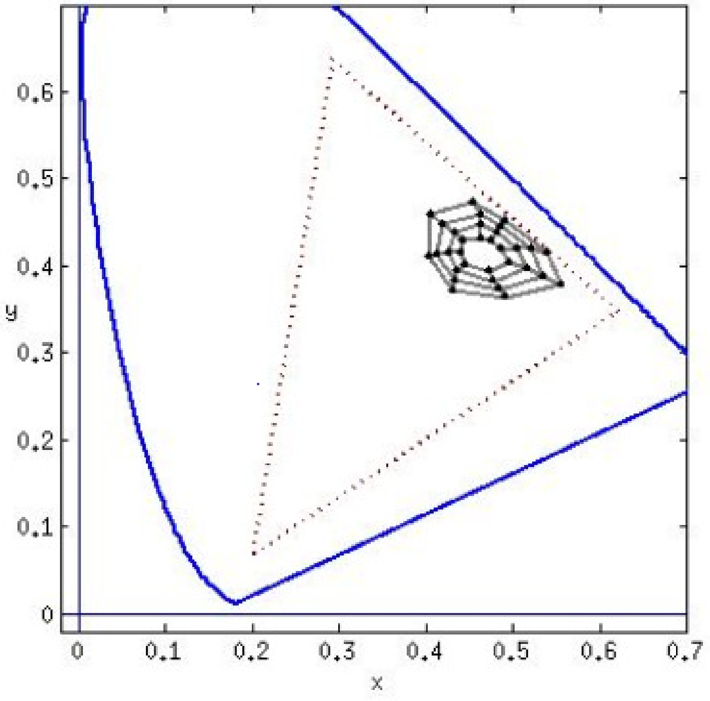


⑤ NEUROSCIENCE FROM STATISTICS (AND VICEVERSA!)

5.1 Neurosci from statistics: COLOR ADAPTATION USING RBIG







⑤ NEUROSCIENCE FROM STATISTICS (AND VICEVERSA!)

5.2 VICEVERSA: Statistics from neuroscience models

TEXTURES

- Analitically: Div. norm. factorizes a plausible PDF model

- Empirically:
  - Bow ties conditional PDFs
  - Mutual information reduction.

⑤ NEUROSCIENCE FROM STATISTICS (AND VICEVERSA!)

5.2 VICEVERSA: Statistics from neuroscience models

TEXTURES

[Malo & Laparra 10]

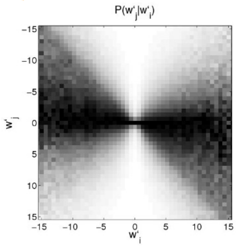
Asuming
- Image model : $P_{w'}(w') \propto "N"(0, \Sigma(w'))$ with $\Sigma(w') = (\beta^\gamma + h|w'|^\gamma)^{\frac{2}{\gamma}}$
- Transform : $r = \dfrac{|w'|^\gamma}{\beta^\gamma + h \cdot |w'|^\gamma}$ →

and using: $P_r(r) = P_{w'}(R^{-1}(r)) \cdot \det(\nabla R^{-1})$

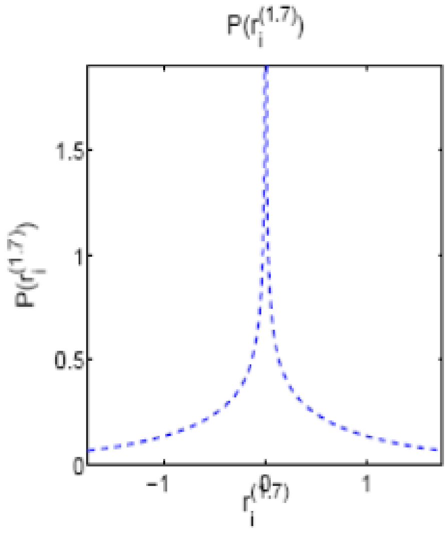
Analytically

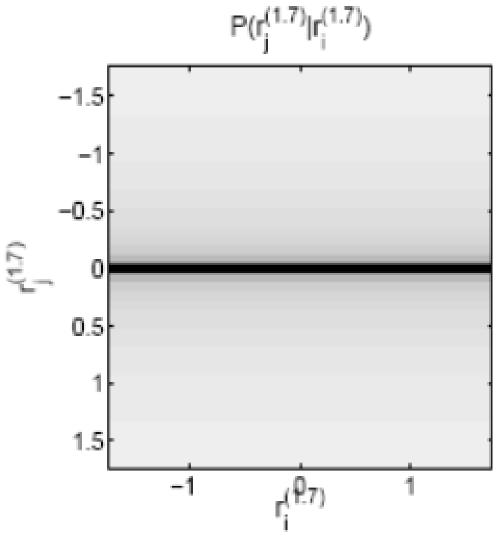
$$\boxed{P_r(r) = \prod_{i=1}^{d} \frac{1}{\gamma (2\pi)^{\frac{1}{2}}} r_i^{\frac{1}{\gamma}-1} e^{-\frac{r_i^{\frac{2}{\gamma}}}{2}} = \prod_{i=1}^{d} P_{r_i}(r_i)}$$

before DIV. NORM. (wavelets)



after DIV. NORM.



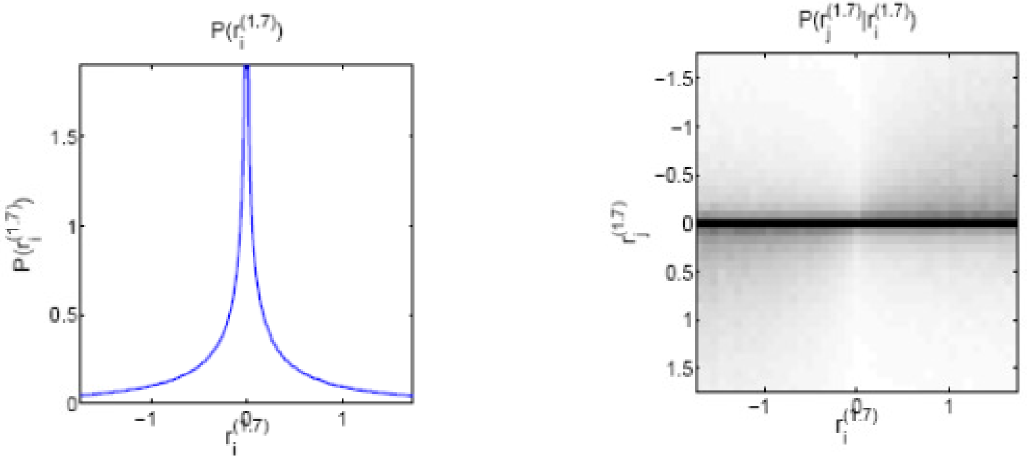


5 NEUROSCIENCE FROM STATISTICS (AND VICEVERSA!)

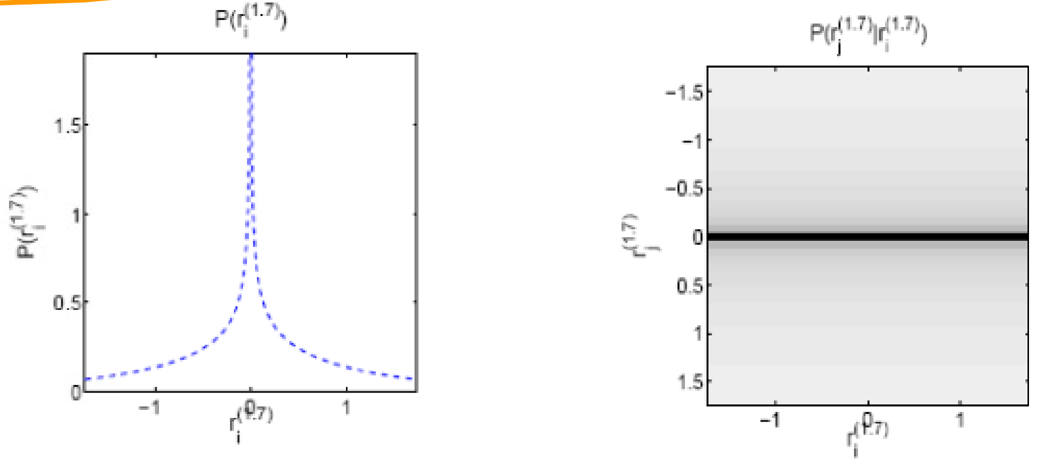
5.2 VICEVERSA: Statistics from neuroscience models TEXTURES

Results: Marginal and conditional PDFs in the divisive normalized domain

EMPIRICAL BEHAVIOR IN DATABASE USING DIV. NORM.



THEORETICAL PREDICTION (DIV. NORM.)

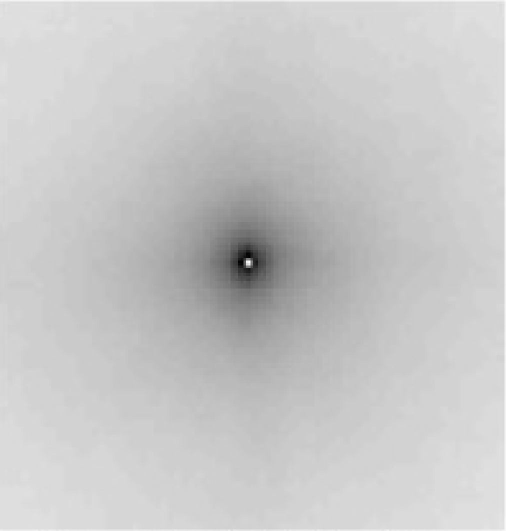


⑤ NEUROSCIENCE FROM STATISTICS (AND VICEVERSA!)

5.2 VICEVERSA: Statistics from neuroscience models   TEXTURES

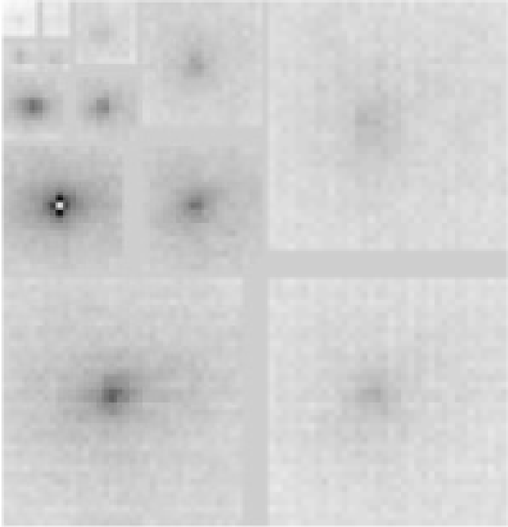
Results: Redundancy reduction (Mutual information in the divisive normalized dom.)

SPATIAL DOMAIN



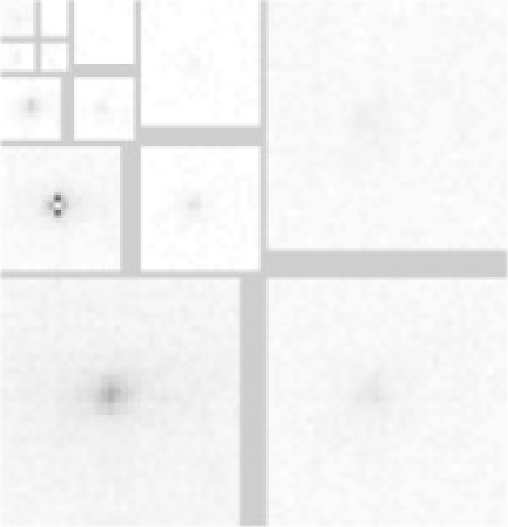
$MI_{\mathbf{x}}$: *0.38*, $[0.21, 1.70]$ bits

(LINEAR)
WAVELETS



$MI_{\mathbf{w}}$: *0.052*, $[10^{-3}, 0.27]$ bits

(NON-LINEAR)
DIVISIVE NORMALIZ.



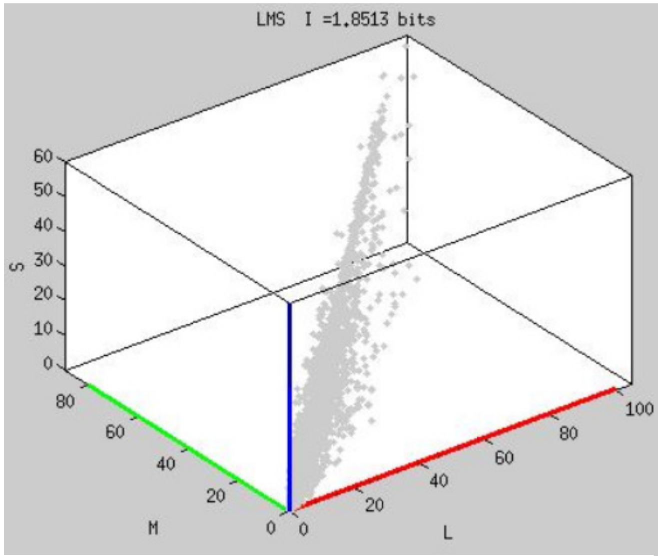
$MI_{\mathbf{r}}$: *0.009*, $[10^{-4}, 0.15]$ bits

⑤ NEUROSCIENCE FROM STATISTICS (AND VICEVERSA!)

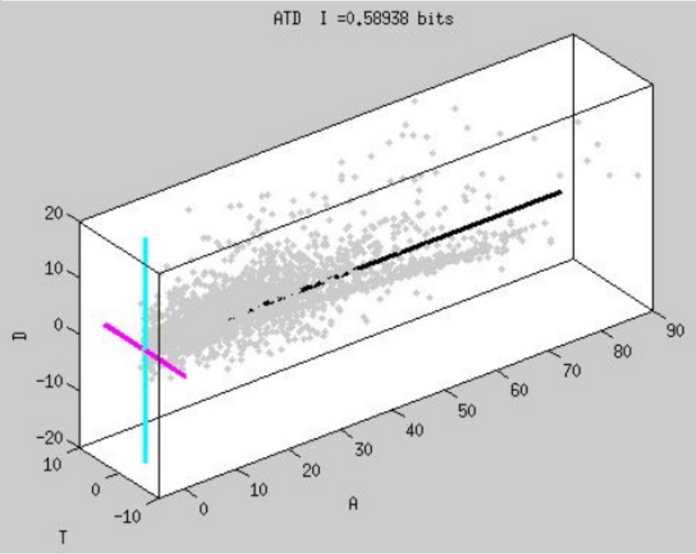
5.2 VICEVERSA: Statistics from neuroscience models COLOR

Empirical color models reduce mutual information.

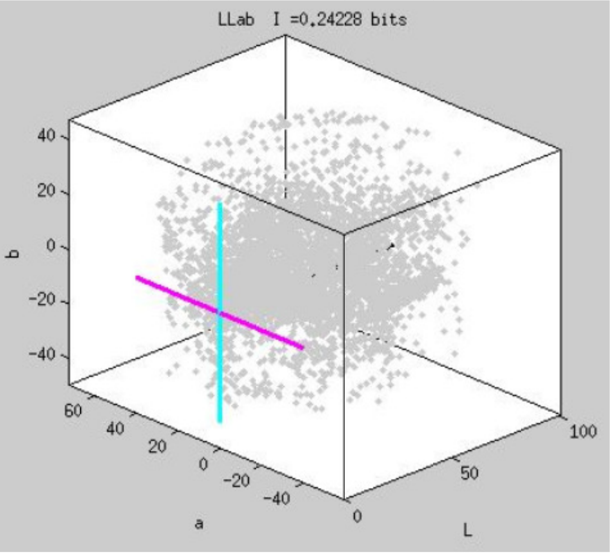
LMS
(I = 1.9 bits)



linear ATD
(I = 0.6 bits)



non-linear LLab
(I = 0.2 bits)



⑤ NEUROSCIENCE FROM STATISTICS (AND VICEVERSA!)

CONCLUSIONS

Your brain seems to be doing a pretty good job (seems Darwin / Barlow were right.)

- * Principled approaches in information theory sense SPCA & RBIG + image & color samples } ⟹ Brain-like behavior (psychophysics)

- * Component independence when applied to natural { colors / images ⟸ Brain-like behavior (empirical models)

FINAL QUESTION: Lets use SPCA, RBIG and Empirical Percept. Models for image processing!

SUMMARY : MANIFOLD LEARNING IN VISION

① Something about me : research agenda THE BARLOW HYPOTHESIS

② Statistical features of images and colors

③ Phenomenology of human vision (in textures and color)

④ Two unsupervised learning algorithms SPCA and RBIG

⑤ Neuroscience from statistics and viceversa

⑥ Applications in image processing

- Coding
- Denoising
- Classification
- Synthesis
- Color constancy
- Image quality

⑥ APPLICATIONS IN IMAGE PROCESSING.

6.1 Image coding

The problem in transform coding:

Find
- statistically
- perceptually

independent representations so that you can quantize each component at a time

Our solution: Use the perception model linear transform + Div. norm since it has good
- statist.
- perceptual

properties

⑥ APPLICATIONS IN IMAGE PROCESSING.

6.1 Image coding

RESULTS [Malo 06, Gomez 05, Camps 08, Patent 08]

⑥ APPLICATIONS IN IMAGE PROCESSING.

6.2 Image denoising

The problems:

$x_n = H \cdot x + n$

- Regularization:
  - $\hat{x} = \arg\min_{x^*} \left(\| H x^* - x_n \| + \lambda \| F(x^*) \| \right)$
  - Kernel regression in transform domain
- Bayesian approach: $\hat{x} = \arg\min_{x^*} \int L(x,x^*) \, P(x^* | x_n) \, dx^*$

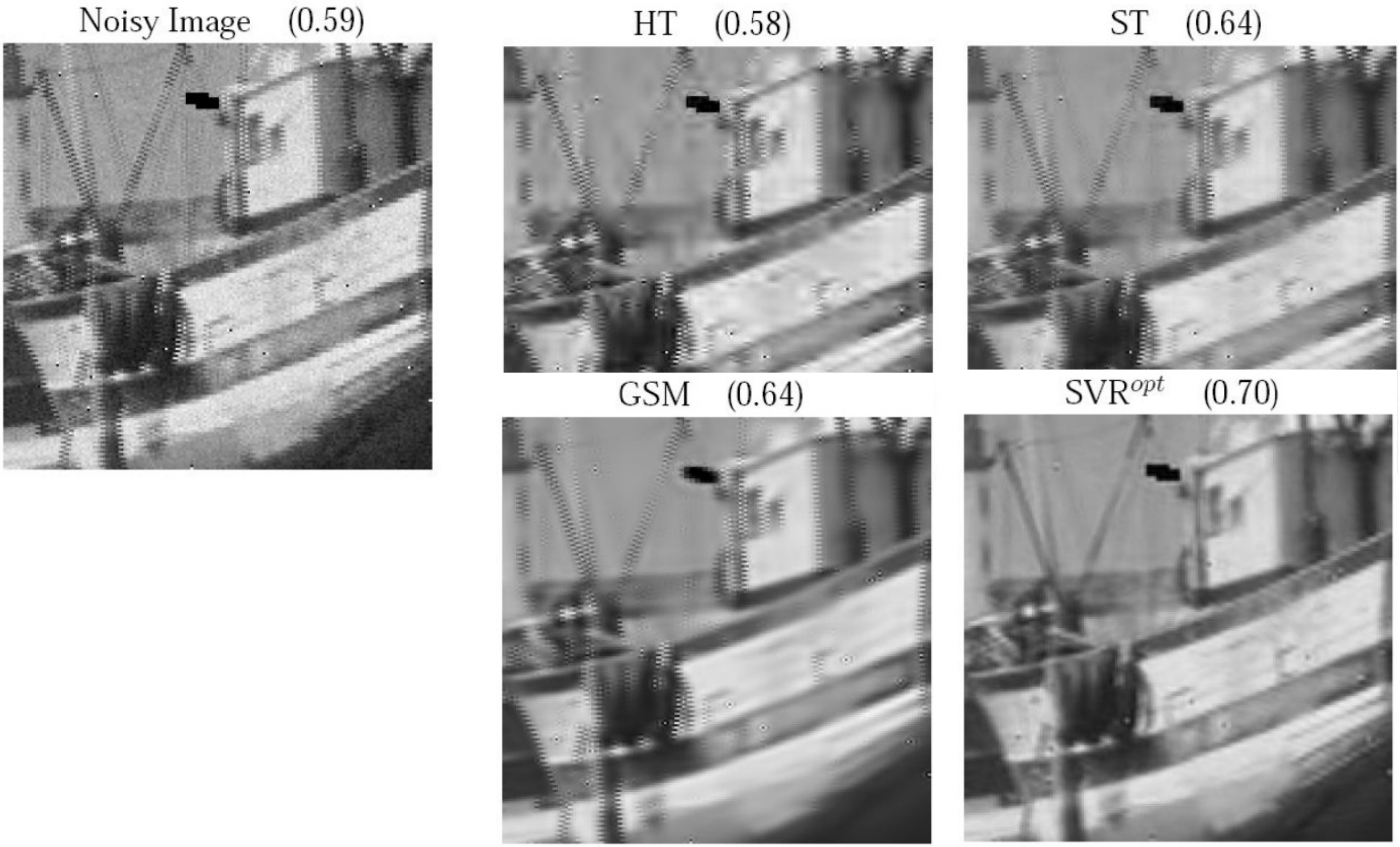
  $P(x^* | x_n) = P(x_n | x^*) \cdot p(x^*)$

Our solutions:

- Regularization
  - Use perceptual penalization functional (Div. Norm.)
  - Design SVM according to image stats.
- Bayesian approach: compute $p(x)$ using RBIG.

⑥ APPLICATIONS IN IMAGE PROCESSING.

6.2 Image denoising RESULTS [Gutierrez 06, Laparra 10a, Laparra 11]



⑥ APPLICATIONS IN IMAGE PROCESSING.

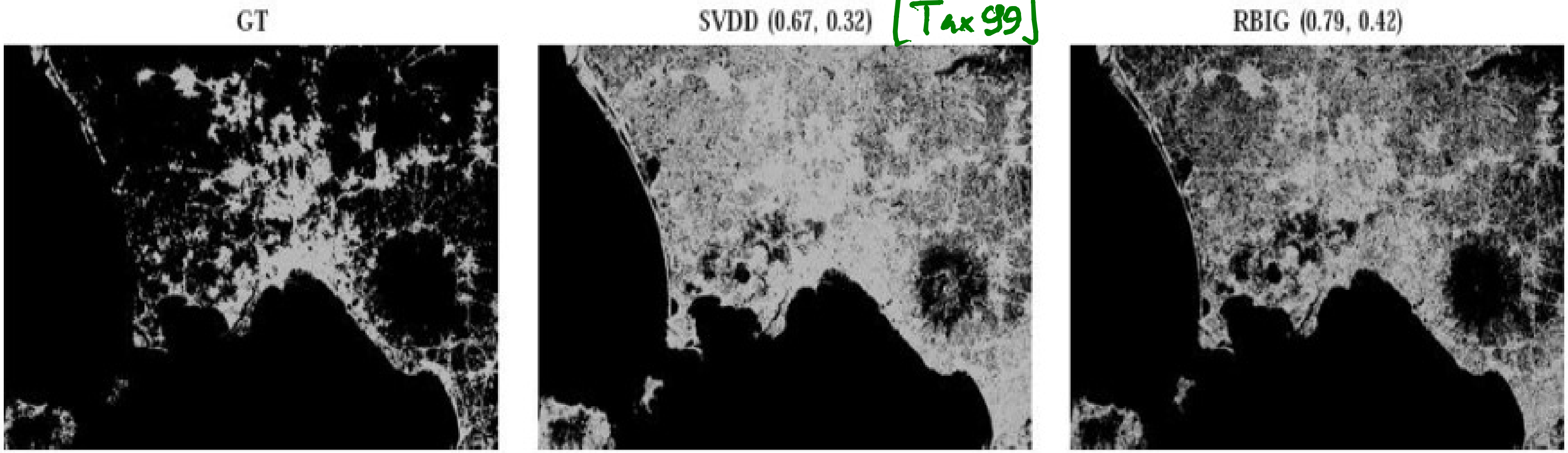
6.3 Image classification

The problem: $x \in C_i$ if $p_{C_i}(x) > p_{C_j}(x) \quad \forall\, C_j$

Our solution: estimate $p_{C_i}(x)$ using RBIG

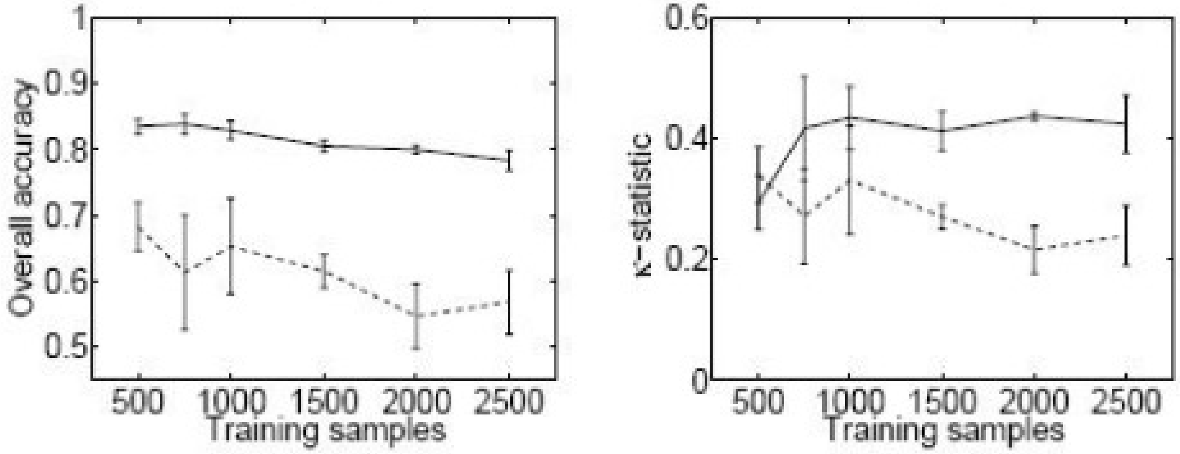
⑥ APPLICATIONS IN IMAGE PROCESSING.

6.3 Image classification: RESULTS [Laparra 11]


GT

SVDD (0.67, 0.32) [Tax 99]

RBIG (0.79, 0.42)



⑥ APPLICATIONS IN IMAGE PROCESSING.

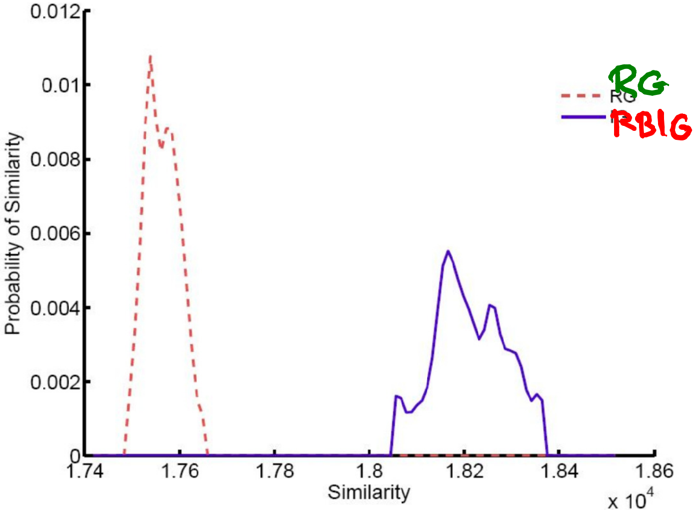
6.4 Image synthesis

The problem: characterize $P(x)$ so you can randomly generate new x samples

Our solution: learn $p(x)$ using RBIG, sample from Gaussian and invert.

⑥ APPLICATIONS IN IMAGE PROCESSING.

6.4 Image synthesis Results [Laparra 11]



⑥ APPLICATIONS IN IMAGE PROCESSING.

6.5 Color constancy (or in Machine Learning the adaptation problem)

The problem: Samples acquired in some observation conditions A and B (e.g. illuminations A and B) follow same PDFs $P_A(x)$ and $P_B(x)$

Given some new sample acquired in conditions B, e.g. x_B, what is the equivalent measurement (e.g. correspond. color) in conditions (illumination) A ?

Our solution: Learn transforms to the canonical domain (using SPCA)

R_A for data set in conditions A

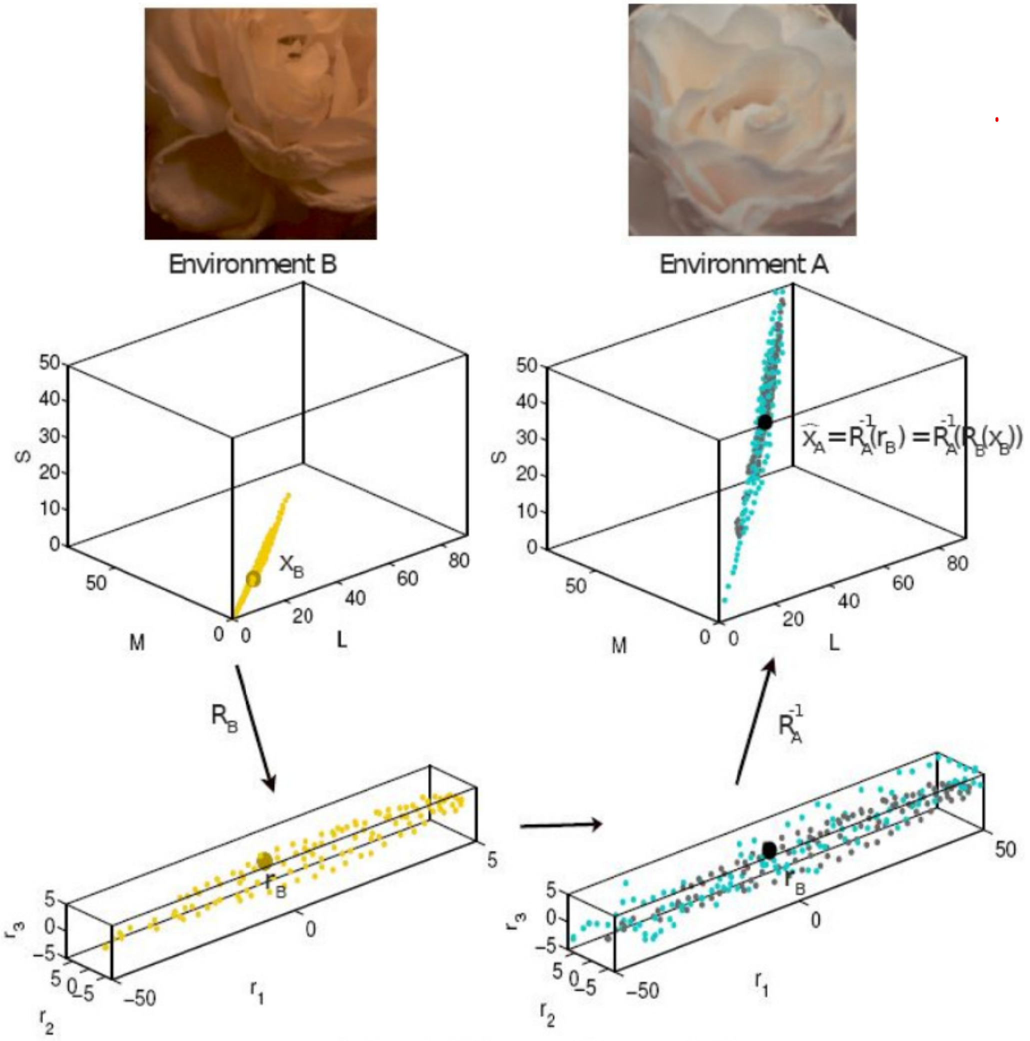
R_B " " " " " B

Then:

$$\hat{x}_A = R_A^{-1}\left(R_B(x_B)\right)$$

⑥ APPLICATIONS IN IMAGE PROCESSING.

6.5 Color constancy RESULTS [Laparra 11, Submitted]



(Canonical) Response Representation

SPCA solution

⑥ APPLICATIONS IN IMAGE PROCESSING.

6.6 Image quality

The problem: Given two images $\left\{ \begin{array}{l} x \\ x+\Delta x \end{array} \right.$ what is the perceptual distance $d_p(x, x+\Delta x)$?

Remember Euclidean metric or $MSE = \|\Delta x\|_2$ is not good!

f_{noise} ~ 3 cpd

f_{noise} ~ 6 cpd

f_{noise} ~ 12 cpd

f_{noise} ~ 24 cpd

Our solution: transform both using div. normaliz. and compute the norm afterwards:

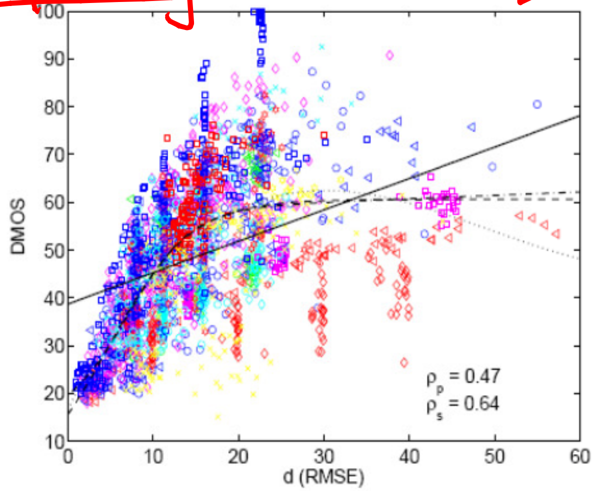
$$x_A = x \longrightarrow r_A$$

$$x_B = x+\Delta x \longrightarrow r_B$$

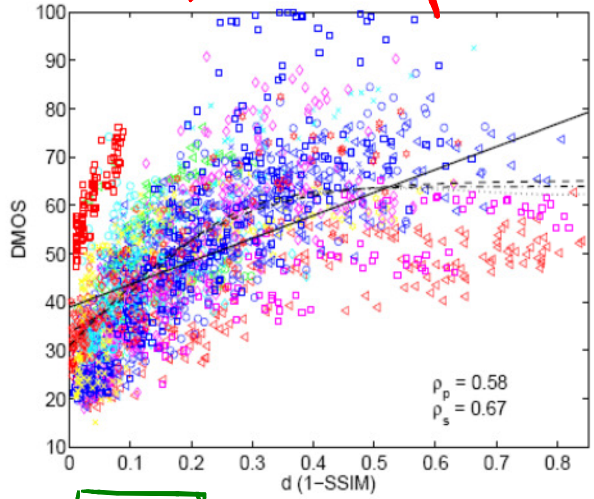
$$d_p(x_A, x_B) = \| r_A - r_B \|_q$$

(6) APPLICATIONS IN IMAGE PROCESSING.

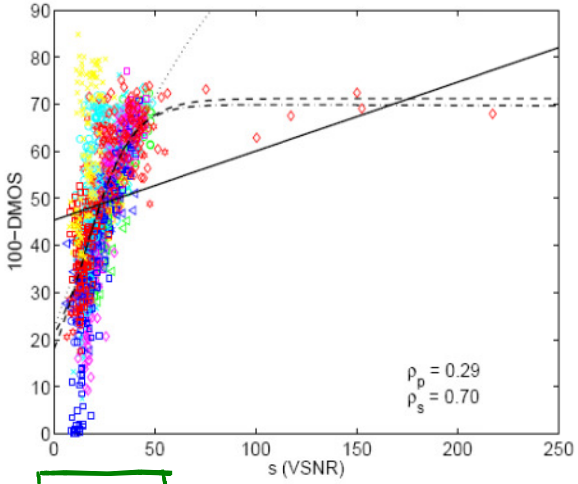
6.6 Image quality: RESULTS [Laparra 10b, Malo & Laparra 10]



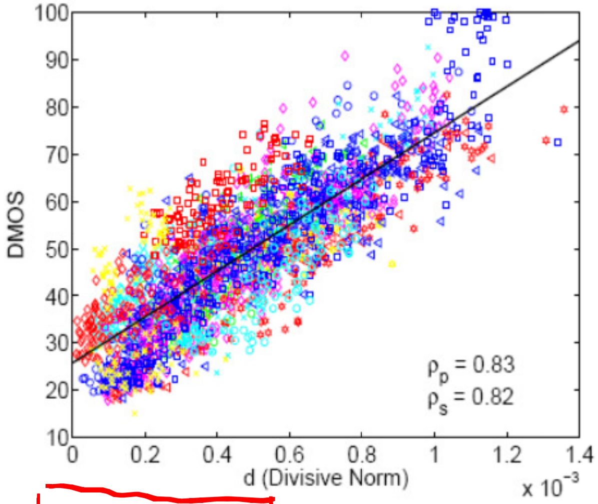
Mean Square Error



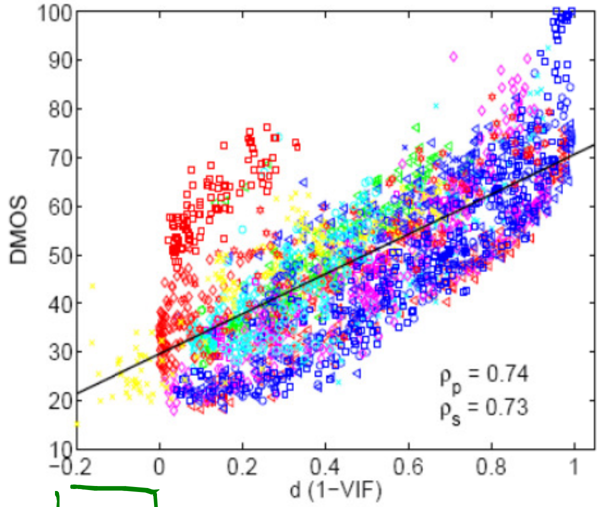
SSIM Wang IEEE TIP 04



VSNR Hemami IEEE TIP 07



Divisiv. Norm. JOSA A 2010



VIF Bovik IEEE TIP 06

MANIFOLD LEARNING IN VISION: CONCLUSIONS

* Spatial patterns and colors in natural images live in very specific manifolds (conventional linear techniques are not enough)

* Your brain
    - · texture sensors
    - · color sensors

  is tuned to specific features and is non-linear

* New unsupervised learning techniques
    - · SPCA
    - · RBIG
    - · Kernels

  allow you to deal with non-linear manifolds

* Your brain may be organized according to
    - · information maximiz.
    - · error minimization

  principles

* This approach
    - - Using empirical perceptual models
    - - Using unsupervised learning

  has lots of applications!