Natural images statistics

Daniela Pamplona

U2IS - ENSTA - IPParis

ecampus moodle: MI210 - Modèles neuro-computationnels de la vision (P4 - 2020-21)

daniela.pamplona@ensta.fr

Contents

Contents

What is vision?

"Vision is the process of discovering from images what is present in the world, and where it is"

David Marr, 1982

What is vision?

"Vision is the process of discovering from images what is present in the world, and where it is"

David Marr, 1982

... in order to solve tasks efficiently.

Daniela Pamplona, 2017

What are the requirements of a working visual system?

Constrains:

- To run real time
- To be robust to noise
- To adapt to lightness, contrast, etc
- To be energetically cheap
- To use limited memory
- To cope with imperfect imaging process

Functions:

- To select and extract relevant information from the environment to solve tasks
- To represent the environment for navigation, reasoning, memory,
- To learn new objects
- To predict location, motion, shape
- …

What are the requirements of a working visual system?

Constrains:

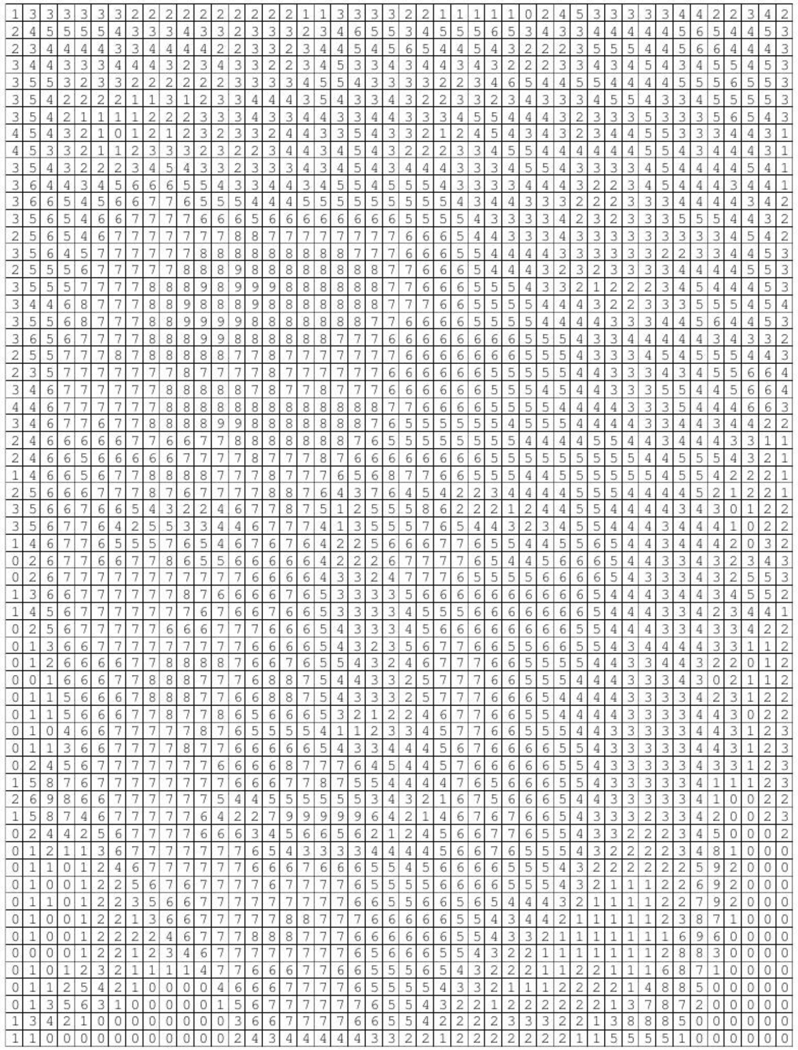
- To run real time
- To be robust to noise
- To adapt to lightness, contrast, etc
- To be energetically cheap
- To use limited ,emory
- To cope with imperfect imaging process

Functions:

- To select and extract relevant information from the environment to solve tasks
- To represent the environment for navigation, reasoning, memory,
- To learn new objects
- To predict location, motion, shape
- ...

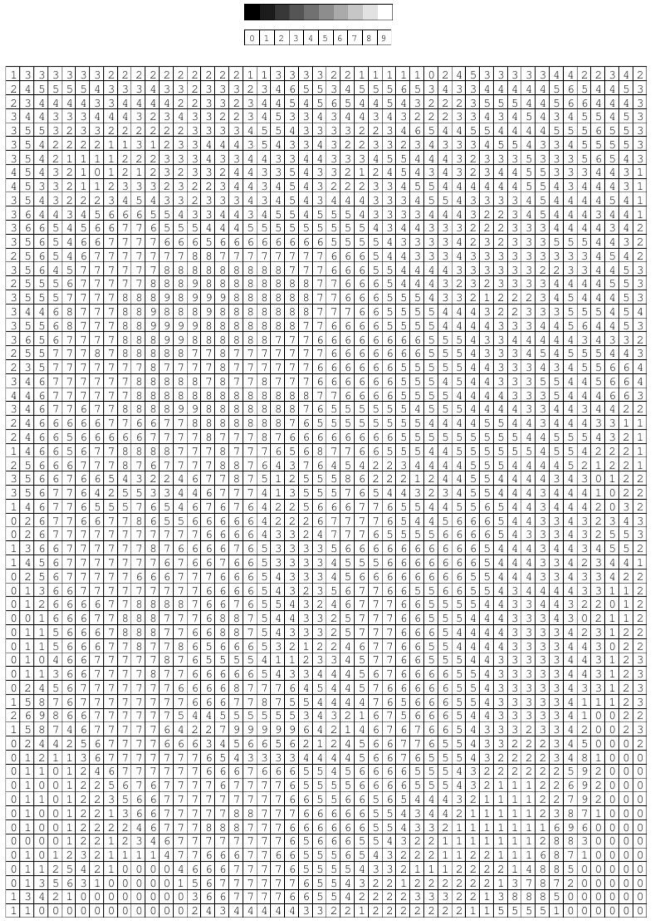
IMPOSSIBLE!

What is in here?



Hyvarinen et al, Natural Image Statistics, 2009

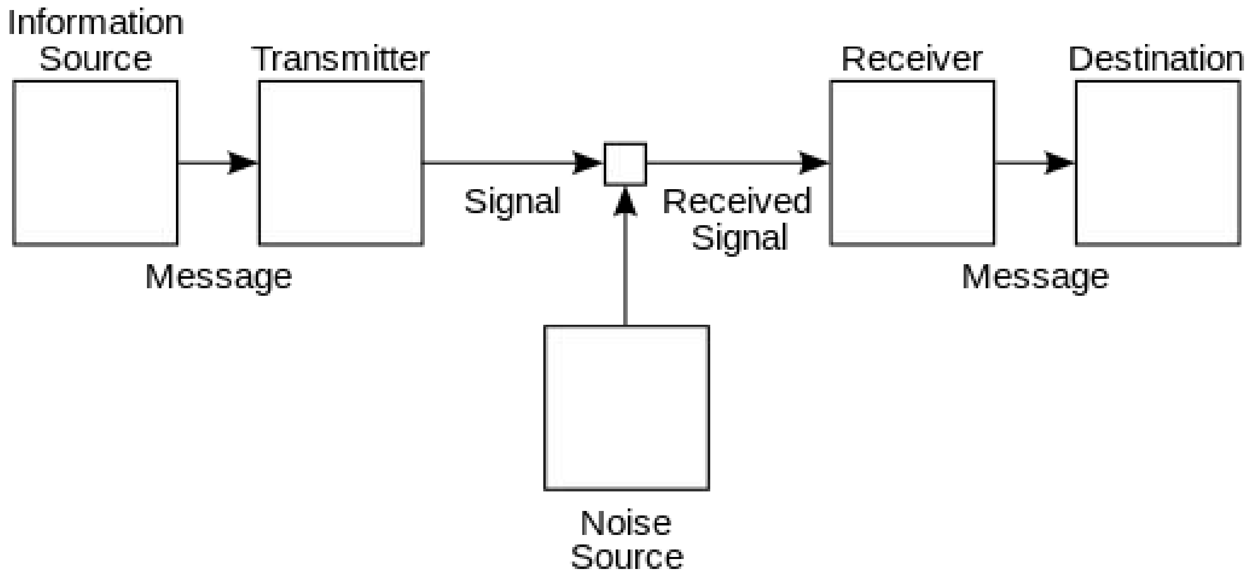
What is in here?



Hyvarinen et al, Natural Image Statistics, 2009

Contents

General Communication System



Definitions

Information Source: Produces messages

Transmitter: Transform/encode the message into a signal

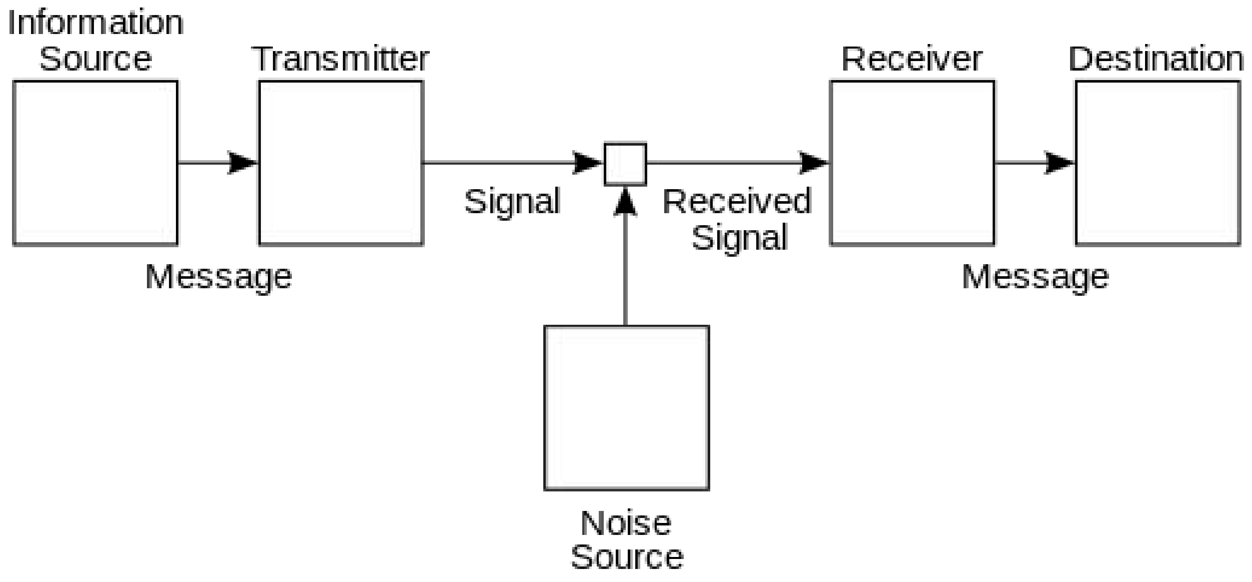
Channel:Medium to carry the message

Receiver: Transforms/decode the signal into a message

Destination: Entity that the message is intended

Shannon A mathematical theory of communication

General Communication System



Example: WhatsApp

Information Source: Alice

Transmitter: Mobile phone

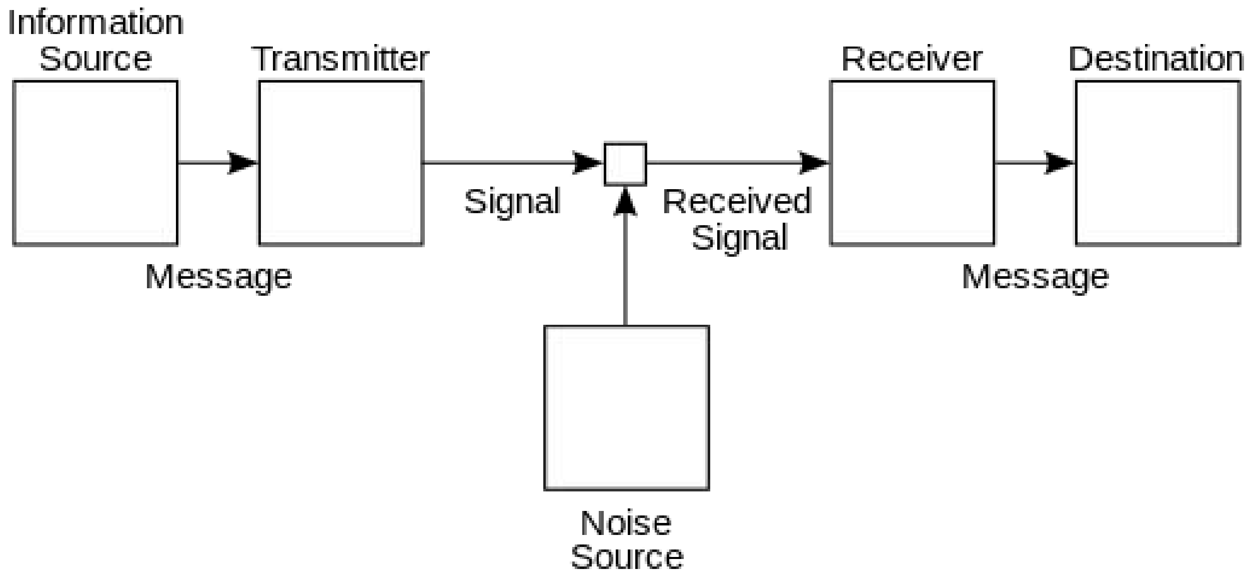
Channel:WiFi

Noise: SNR:30dB

Receiver: Mobile phone

Destination: Bob

General Communication System



Example: Visual System

Information Source: Environment

Transmitter: Eye

Channel: Early visual system

Noise: Unknown

Receiver: Higher areas (MT,TE,MIP,..)

Destination: Other brain areas (PMC,..)
(ultimatly the environment)

Self-Information of a message

Information content of an event, $h(E)$, with probability $P\{E\}$ must:

1. be a decreasing function of $P\{E\}$: more an event is likely, the less information its occurrence brings to us.

Self-Information of a message

Information content of an event, $h(E)$, with probability $P\{E\}$ must:

1. be a decreasing function of $P\{E\}$: more an event is likely, the less information its occurrence brings to us.
2. if $P\{E\} = 1$ then $h(E) = 0$, if we are certain that E will occur, we get no information from its outcome.

Self-Information of a message

Information content of an event, $h(E)$, with probability $P\{E\}$ must:

1. be a decreasing function of $P\{E\}$: more an event is likely, the less information its occurrence brings to us.
2. if $P\{E\} = 1$ then $h(E) = 0$, if we are certain that E will occur, we get no information from its outcome.
3. if E and F are independent, $h(E \text{ and } F) = h(E) + h(F)$

Self-Information of a message

Information content of an event, $h(E)$, with probability $P\{E\}$ must:

1. be a decreasing function of $P\{E\}$: more an event is likely, the less information its occurrence brings to us.
2. if $P\{E\} = 1$ then $h(E) = 0$, if we are certain that E will occur, we get no information from its outcome.
3. if E and F are independent, $h(E \text{ and } F) = h(E) + h(F)$

$$h(E) = \log \frac{1}{P\{E\}} = -\log P\{E\}$$

Entropy

- Entropy: measure of randomness of a random variable (r. v.)
- Entropy is the expected value of self information

Entropy

- Entropy: measure of randomness of a random variable (r. v.)
- Entropy is the expected value of self information

X : discrete r.v. taking values in $\{x_1, x_2, \ldots, x_n\}$ with $p_i = P\{X = x_i\}$

$$H(X) = -\sum_{i=1}^{n} p_i \log p_i$$

Maximum entropy of a discrete r.v.

- X: Bernoulli r. v.,

  $H_{max}(X) = 1$ with $p = 0.5$

- X : discrete r.v. taking values in $\{ x_1 , x_2 ,..., x_n \}$ with $p_i = 1/n$

  $H_{max}(X) = Log(N)$

Redundancy of a r.v.

- X: Bernoulli r. v.,
  $H_{max}(X) = 1$ with $p = 0.5$

- X : discrete r.v. taking values in $\{x_1, x_2, \ldots, x_n\}$ with $p_i = 1/n$
  $H_{max}(X) = Log(N)$

- Redundancy, or relative entropy, compares the entropy of a random variable with the maximal entropy
  $r = 1 - H(X)/H_{max}(X)$

Redundancy of a r.v.

- X: Bernoulli r. v.,

  $H_{max}(X) = 1$ with $p = 0.5$

- X : discrete r.v. taking values in $\{x_1, x_2, \ldots, x_n\}$ with $p_i = 1/n$

  $H_{max}(X) = Log(N)$

- Redundancy, or relative entropy, compares the entropy of a random variable with the maximal entropy

  $r = 1 - H(X)/H_{max}(X)$

REDUCE THE REDUNDANCY <=> INCREASE THE ENTROPY

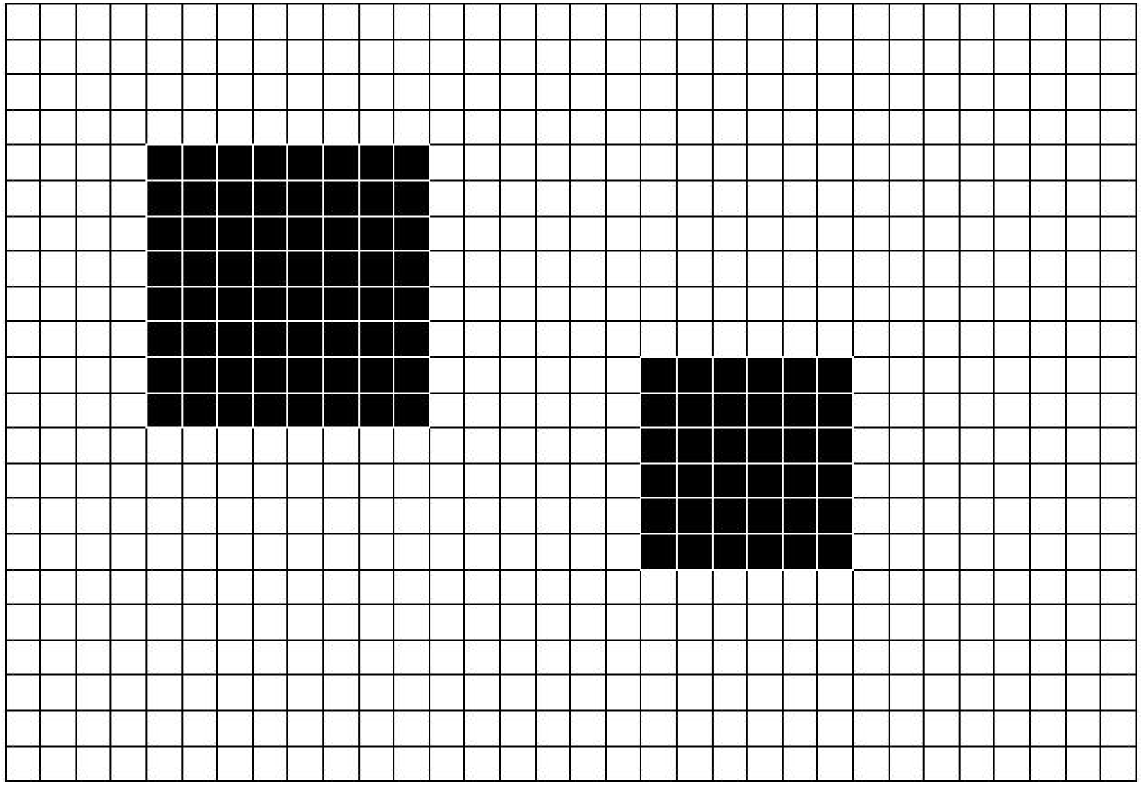
Contents

What is the function of the early visual system?

They [The sensory relays] recode sensory messages, extracting signals of high relative entropy from the highly redundant sensory input.

Barlow 1961

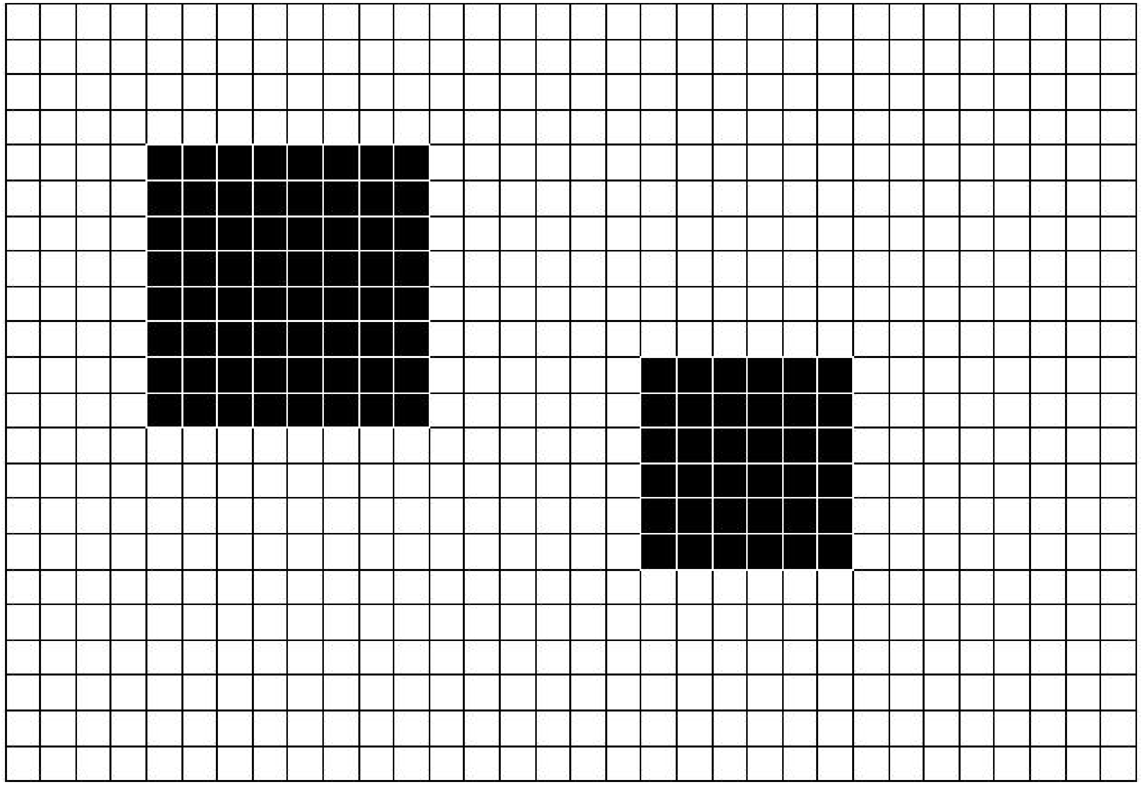
Redundancy



What is present in this world and where is it?

Attneave, Some informational aspects of visual perception

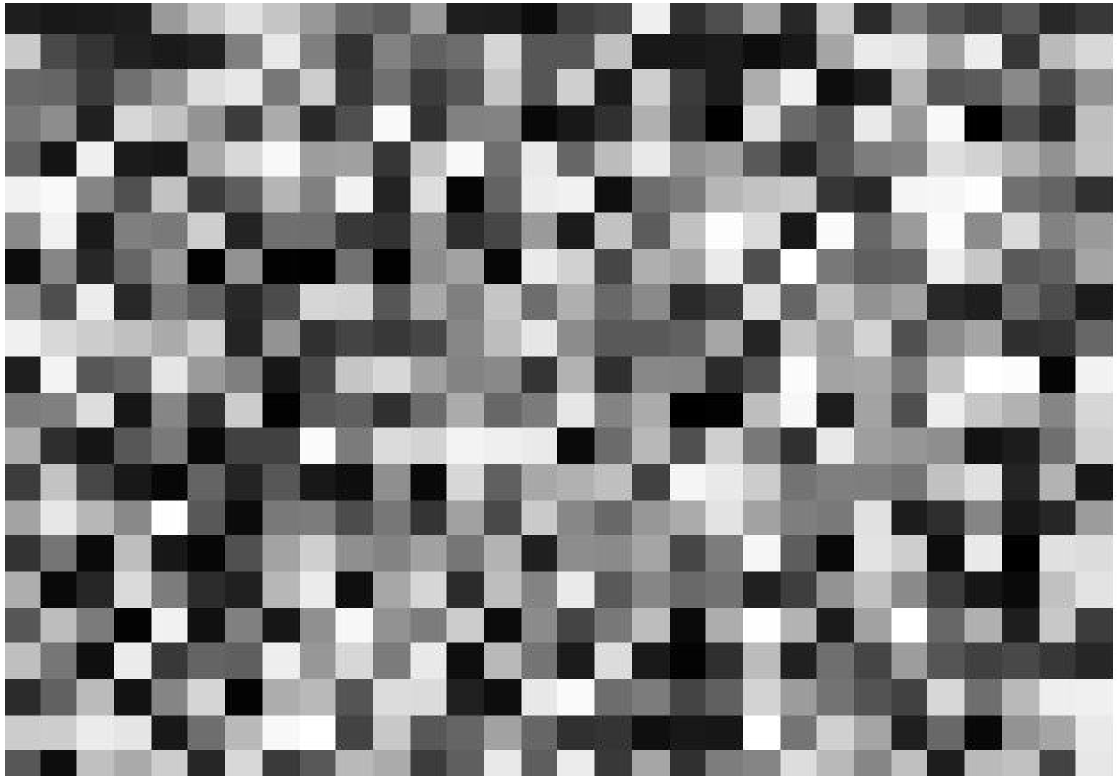
Redundancy



What is present in this world and where is it?

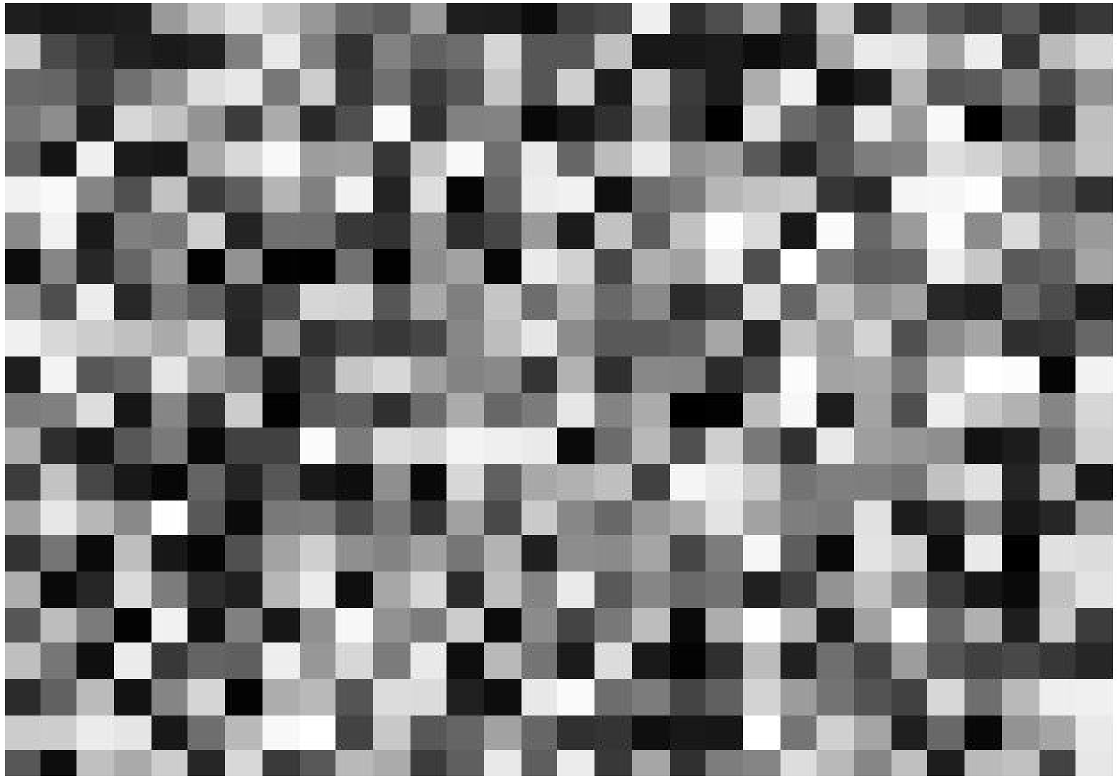
Black square, 8, (4,4)
Black square, 6, (18,14)

Redundancy



What is present in this world and where is it?

Redundancy



What is present in this world and where is it?

Black pixel (0,0)
Black pixel (0,1)
Black pixel (0,2)
Black pixel (0,3)
Gray pixel (0,4)
Light gray pixel (0,4)
...

Redundancy

What is present in this world and where is it?

Redundancy

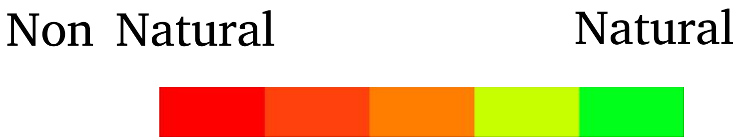
What is present in this world and where is it?

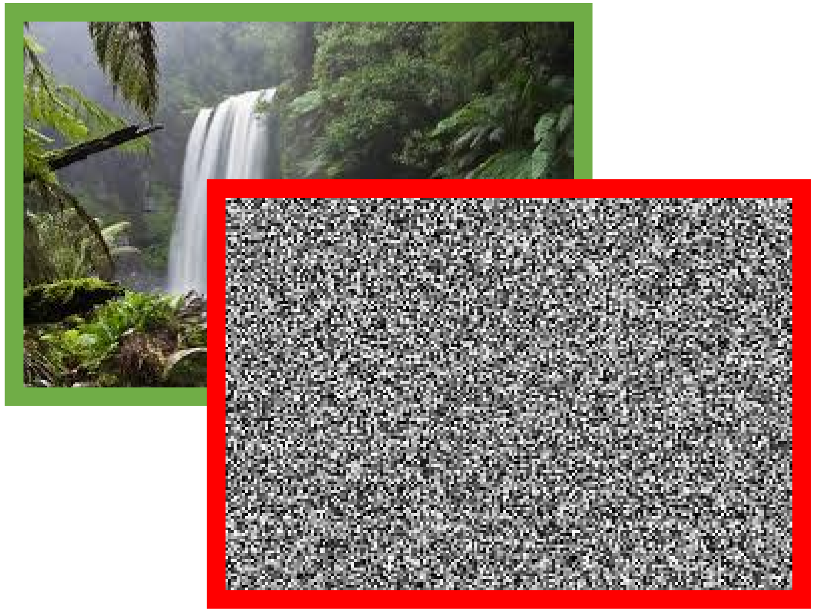
Branches and leaves on top
Mountain on left to center
Reflects on left to center bottom
Dog and person on bottom
Dense trees on right
...

What is the message?
What are natural images?

Non Natural Natural

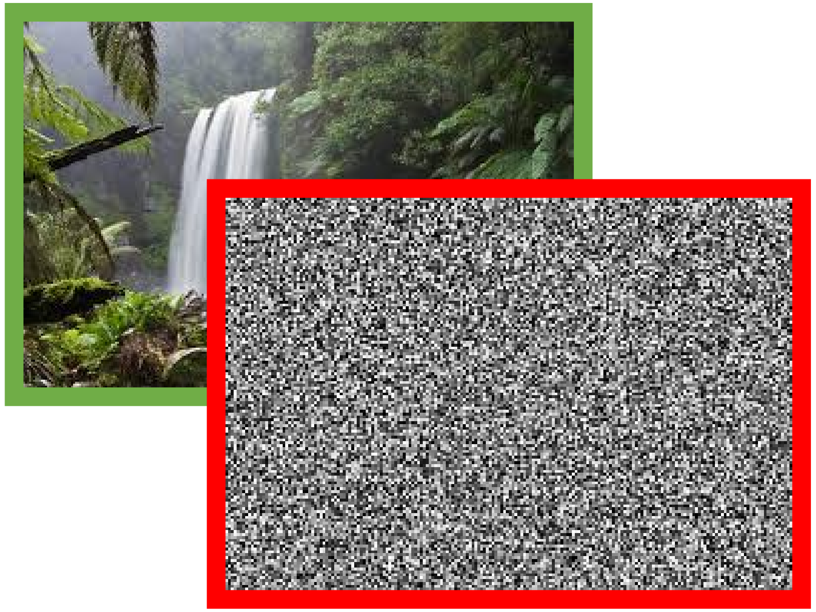
What is the message?
What are natural images?





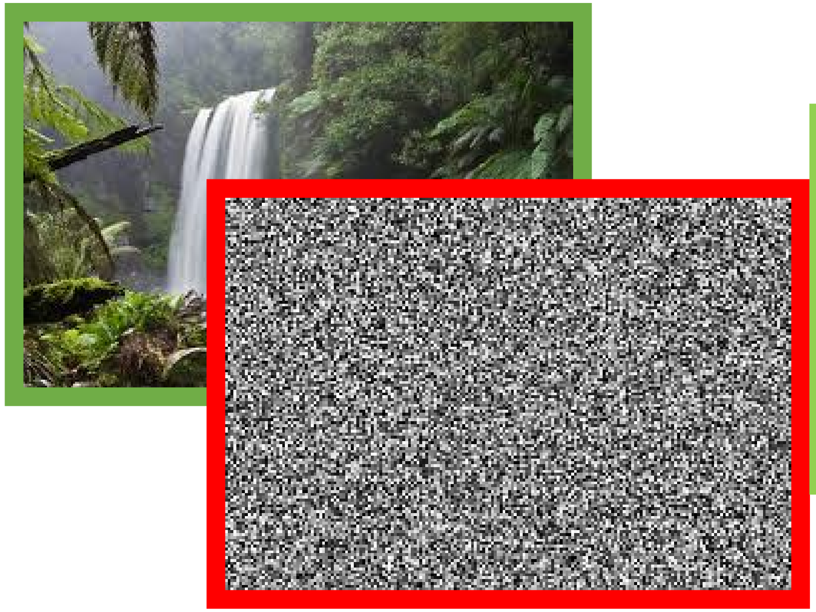
What is the message?
What are natural images?

Non Natural Natural



What is the message?
What are natural images?

Non Natural Natural



What is the message?
What are natural images?

Non Natural Natural

What is the message?
What are natural images?

Non Natural Natural

What is the message?
What are natural images?

Non Natural — Natural

What is the message?
What are natural images?

Question: Can we measure the redundancy of natural images?

Question: Can we measure the redundancy of natural images?

Answer: No, because we do not know the probability distribution of natural images
(we do know the distribution of white noise images, e.g. $N(0,1)$)

Question: Can we measure the redundancy of natural images?

Answer: No, because we do not know the probability distribution of natural images
(we do know the distribution of white noise images, e.g. $N(0,1)$)

Question: Can we, at least, approximate it?

Question: Can we measure the redundancy of natural images?

Answer: No, because we do not know the probability distribution of natural images
(we do know the distribution of white noise images, e.g. N(0,1))

Question: Can we, at least, approximate it?

Answer: Yes!

Question: Can we measure the redundancy of natural images?

Answer: No, because we do not know the probability distribution of natural images
(we do know the distribution of white noise images, e.g. $N(0,1)$)

Question: Can we, at least, approximate it?

Answer: Yes!

Question: How?

Question: Can we measure the redundancy of natural images?

Answer: No, because we do not know the probability distribution of natural images
(we do know the distribution of white noise images, e.g. $N(0,1)$)

Question: Can we, at least, approximate it?

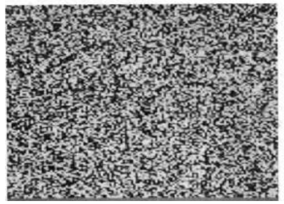
Answer: Yes!

Question: How?

Answer: Looking at the statistics of the natural images

Natural images are redundant

100% deleted

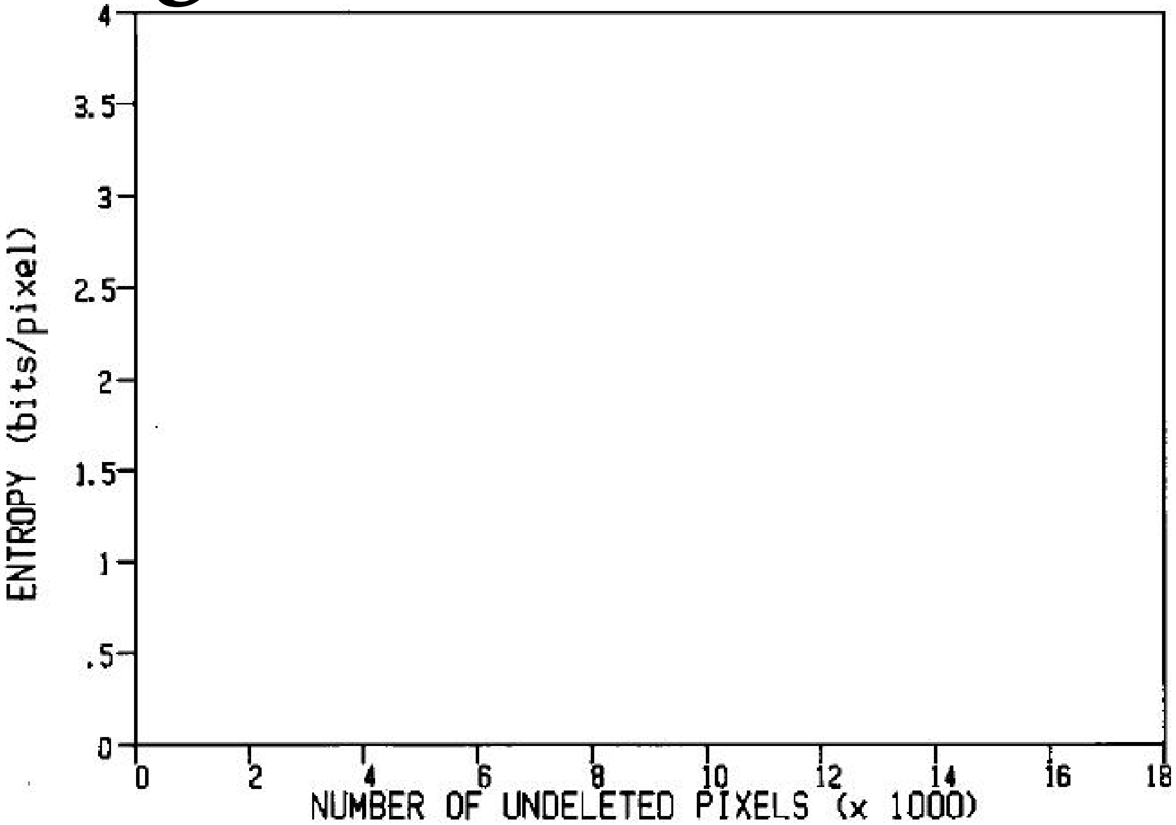


40% deleted

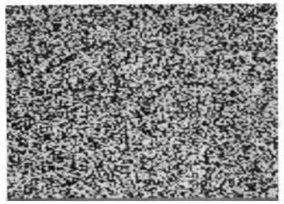
1% deleted

Kersten Predictability and redundancy of natural images,

Natural images are redundant



100% deleted

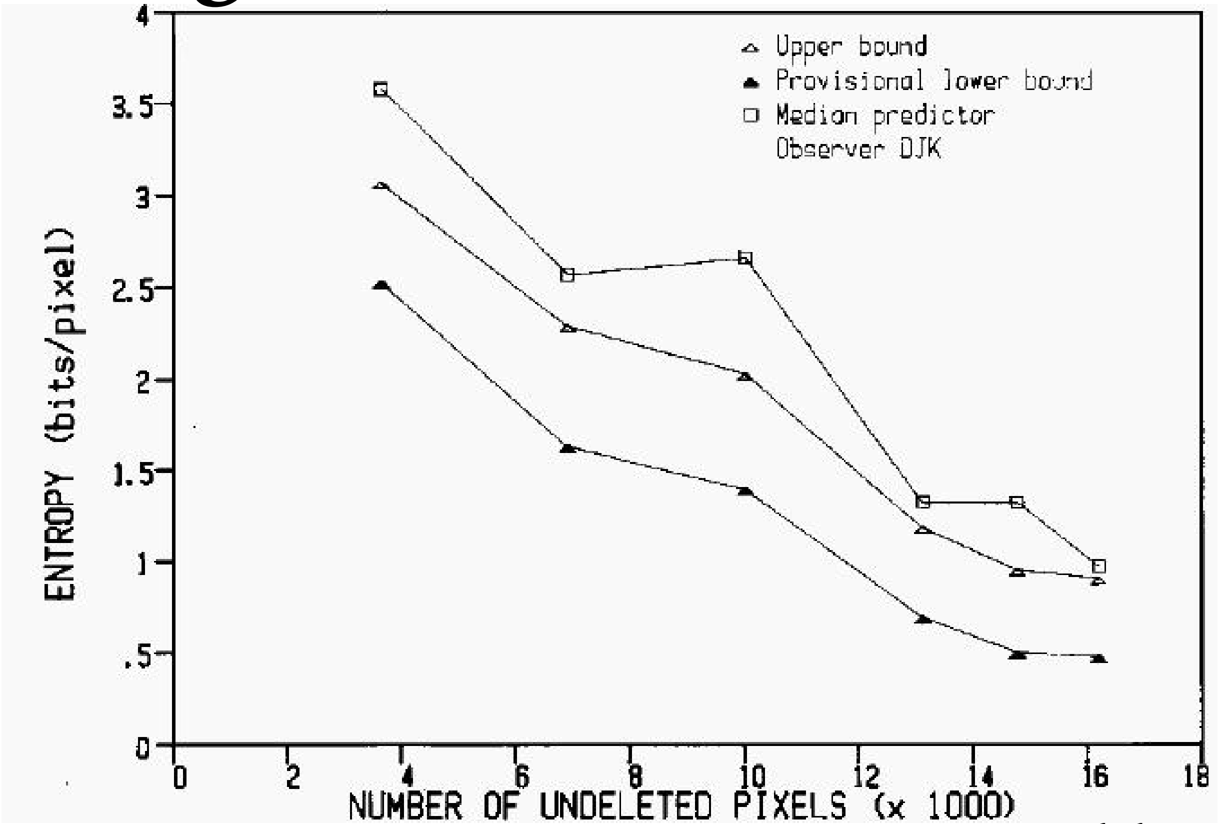


40% deleted

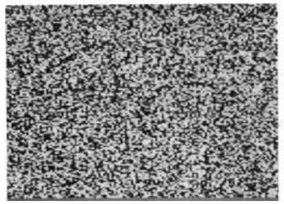
1% deleted

Kersten Predictability and redundancy of natural images,

Natural images are redundant



100% deleted

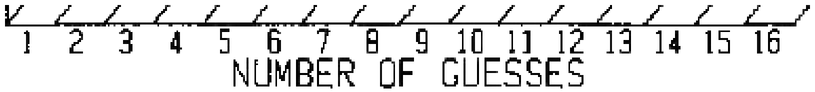


40% deleted

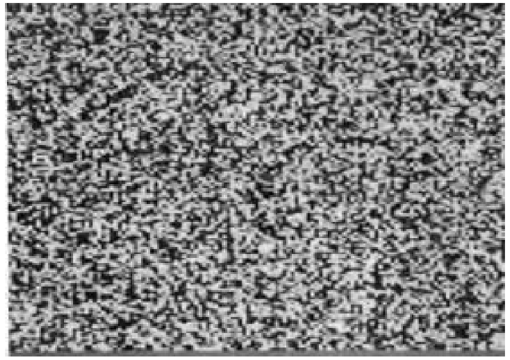
1% deleted

Kersten Predictability and redundancy of natural images,

Natural images are redundant. So what?



100% deleted

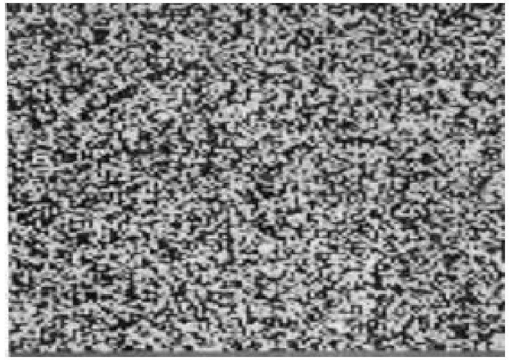


16 gray levels

Kersten Predictability and redundancy of natural images,

Natural images are redundant. So what?

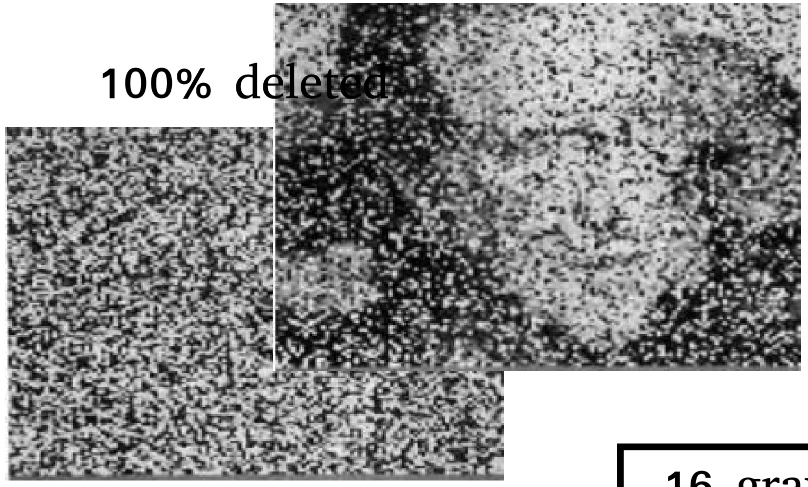
100% deleted



16 gray levels

Kersten Predictability and redundancy of natural images,

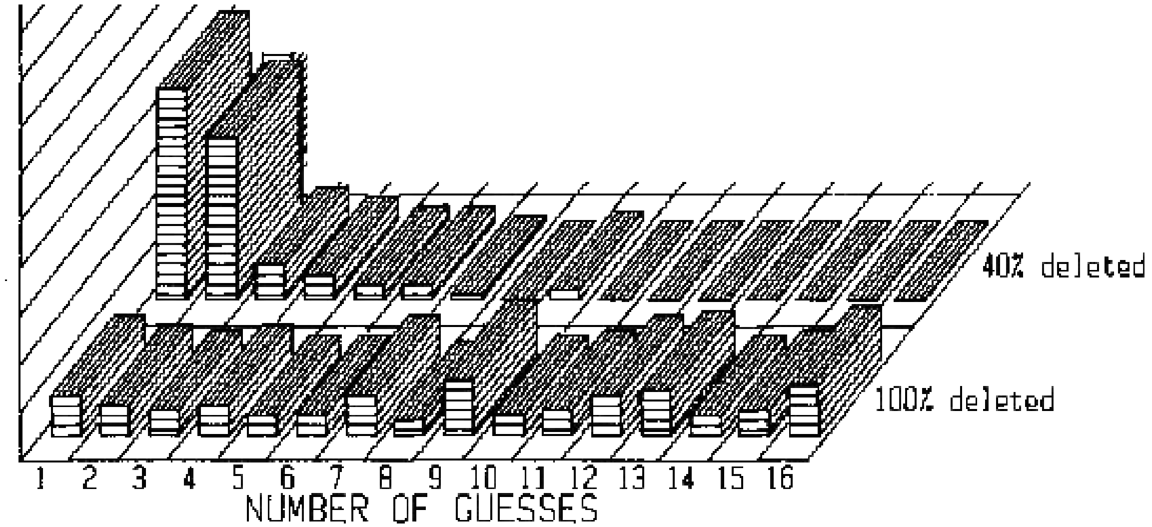
Natural images are redundant. So what?

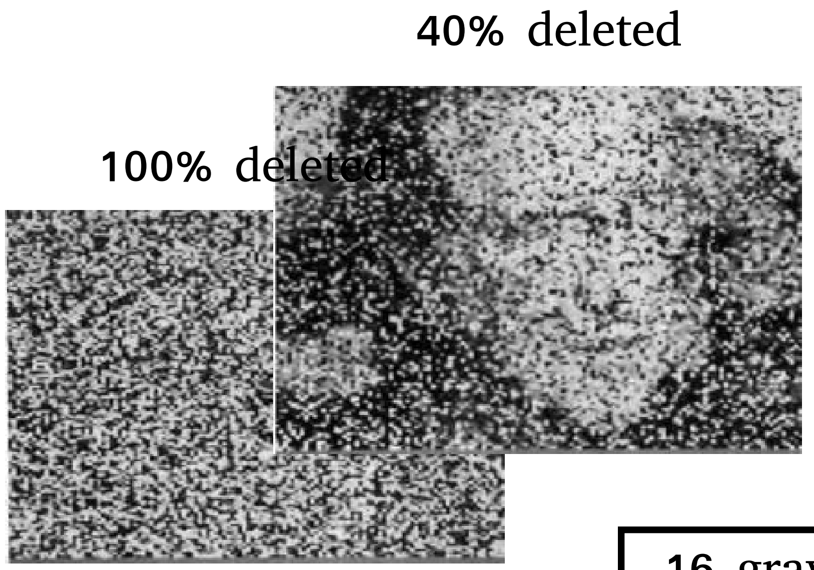


16 gray levels

Kersten Predictability and redundancy of natural images,

Natural images are redundant. So what?

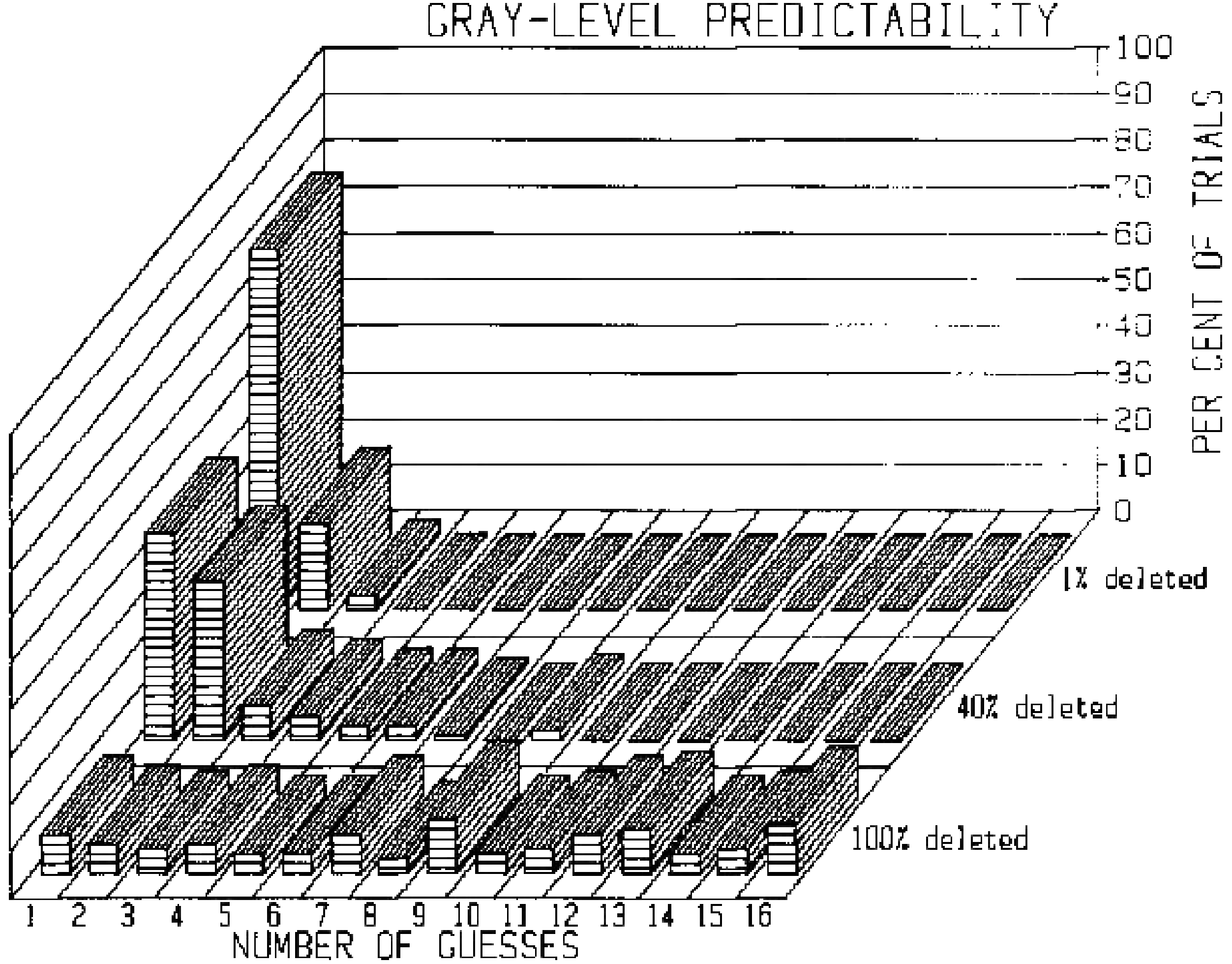




16 gray levels

Kersten Predictability and redundancy of natural images,

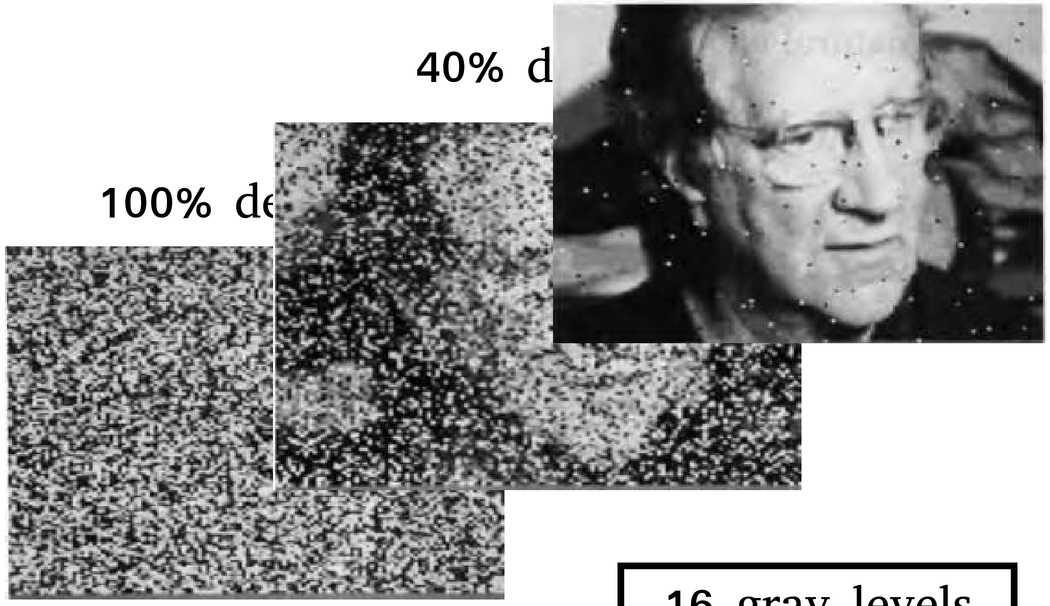
Natural images are redundant. So what?



1% deleted

40% d

100% de



16 gray levels

Kersten Predictability and redundancy of natural images,

Contents

General idea (method of the moments):

- μ^n moment of order n: $\mu^n[X] = E[(X-\mu)^n]$
- X r. v. with n - first moments well defined
- $X_1, X_2, \ldots, X_k$ sequence of r. v.

If $\lim_{k\to\infty} \mu^n[X^k] = \mu^n[X]$ then $X^k \xrightarrow{d} X$

Moments

- $\mu^1[X] = E[X-\mu]=0$
- $\mu^2[X] = E[(X-\mu)^2]$ (auto correlation)
- $\mu^3[X] = E[(X-\mu)^3]$ (skewness)
- $\mu^4[X] = E[(X-\mu)^4]$ (kurtosis)
- ...

Fourier Transform

- The Fourier transform decomposes a function of time (a signal) into the frequencies that make it up.

1D

$$F(x) = \sum_{n=0}^{N-1} f(n) e^{-j2\pi(x\frac{n}{N})}$$

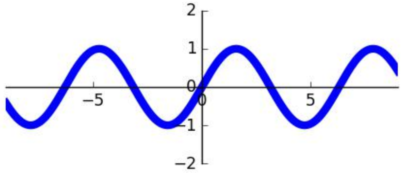
$$f(n) = \frac{1}{N} \sum_{n=0}^{N-1} F(x) e^{j2\pi(x\frac{n}{N})}$$

$$PS(x) = |F(x)|^2$$

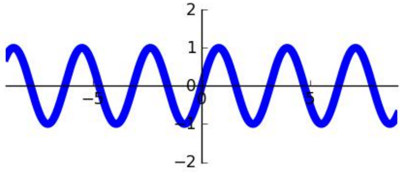
1D Fourier Transform

time domain

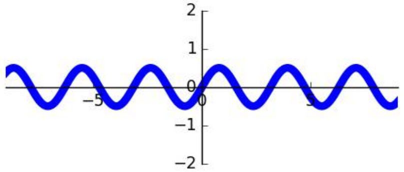
sin(x)



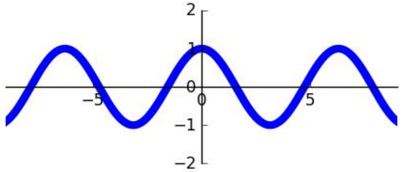
sin(2x)



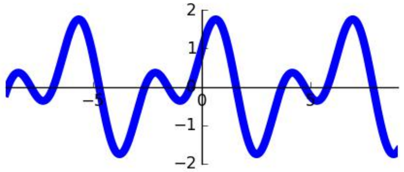
0.5sin(2x)



sin(x+pi/2)



sin(2x)+sin(x+pi/2)

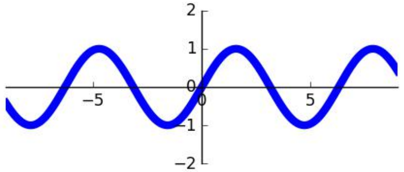


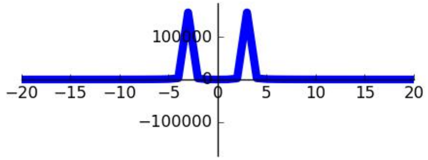
1D Fourier Transform

time domain

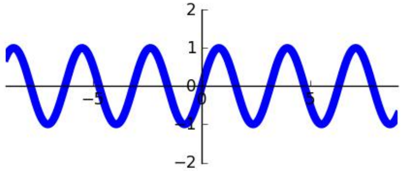
ps (fft)

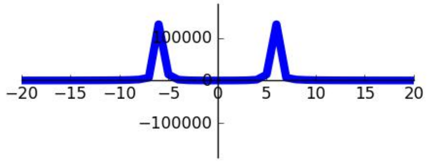
sin(x)



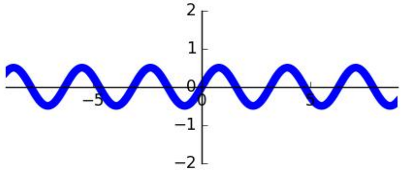


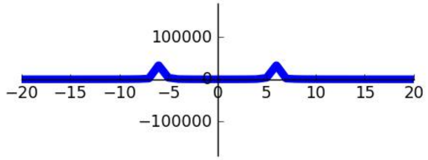
sin(2x)



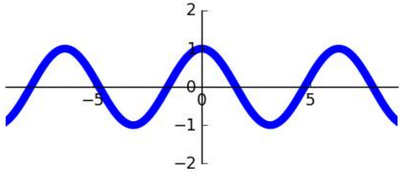


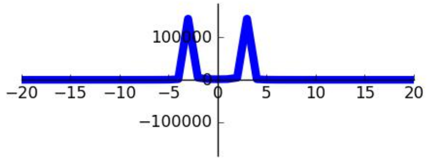
0.5sin(2x)



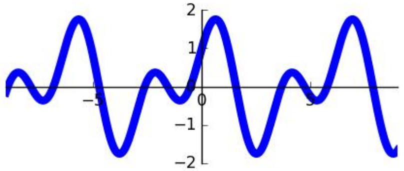


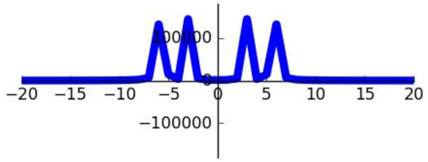
sin(x+pi/2)





sin(2x)+sin(x+pi/2)





Fourier Transform

- The Fourier transform decomposes a function of time (a signal) into the frequencies that make it up.

1D

$$F(x) = \sum_{n=0}^{N-1} f(n) e^{-j2\pi(x\frac{n}{N})}$$

$$f(n) = \frac{1}{N}\sum_{n=0}^{N-1} F(x) e^{j2\pi(x\frac{n}{N})}$$

$$PS(x) = |F(x)|^2$$

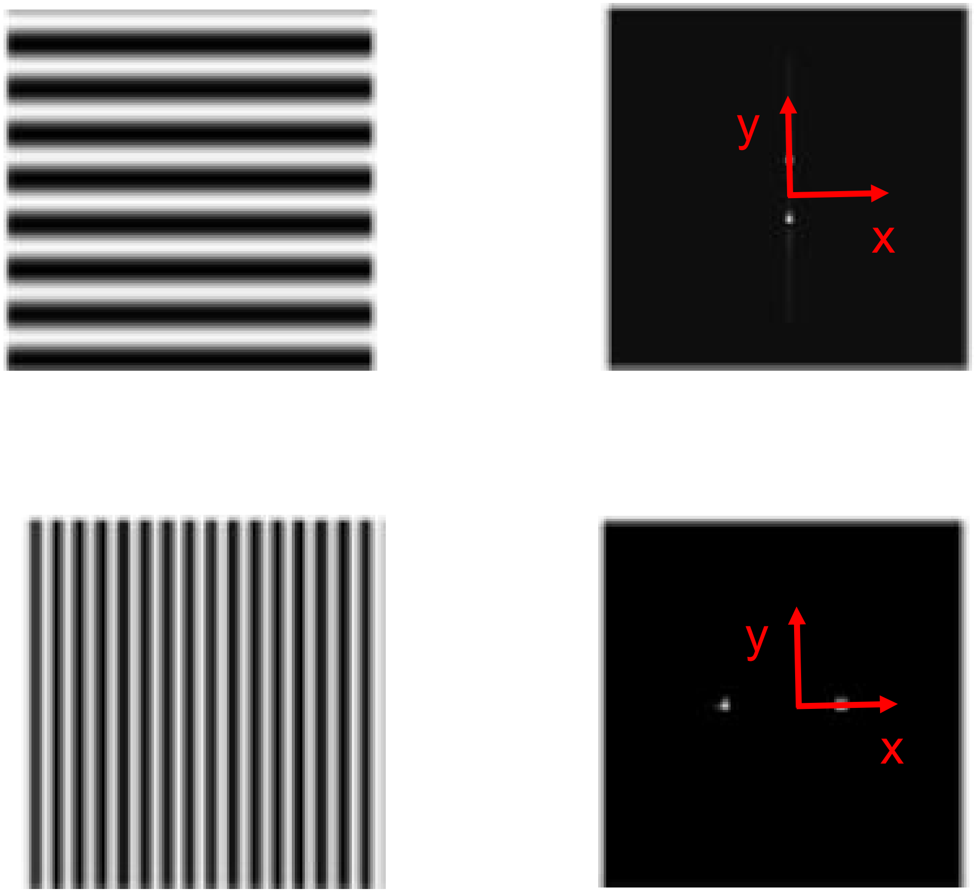
2D

$$F(x,y) = \sum_{m=0}^{M-1}\sum_{n=0}^{N-1} f(m,n) e^{-j2\pi(x\frac{m}{M}+y\frac{n}{N})}$$

$$f(m,n) = \frac{1}{MN}\sum_{m=0}^{M-1}\sum_{n=0}^{N-1} F(x,y) e^{j2\pi(x\frac{m}{M}+y\frac{n}{N})}$$

$$PS(x,y) = |F(x,y)|^2$$

2 D Fourier Transform of waves



Fourier Transform

- The Fourier transform decomposes a function of time (a signal) into the frequencies that make it up.

1D

$$F(x) = \sum_{n=0}^{N-1} f(n) e^{-j2\pi(x\frac{n}{N})}$$

$$f(n) = \frac{1}{N} \sum_{n=0}^{N-1} F(x) e^{j2\pi(x\frac{n}{N})}$$

$$PS(x) = |F(x)|^2$$

2D

$$F(x,y) = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} f(m,n) e^{-j2\pi(x\frac{m}{M}+y\frac{n}{N})}$$

$$f(m,n) = \frac{1}{MN} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} F(x,y) e^{j2\pi(x\frac{m}{M}+y\frac{n}{N})}$$

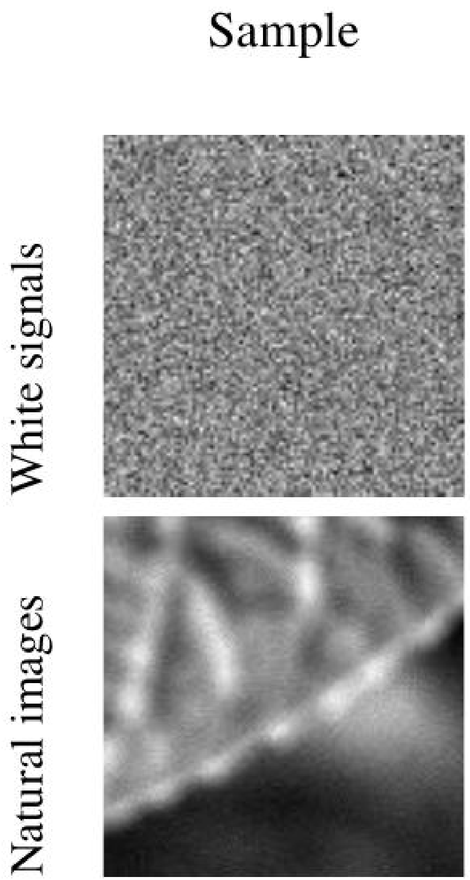
$$PS(x,y) = |F(x,y)|^2$$

Wiener-Khinchin theorem: If $X = \{ x\_1, \ldots, x\_n \}$ is a stationary process and $\mu^2[X]$ exists and is finite, then:

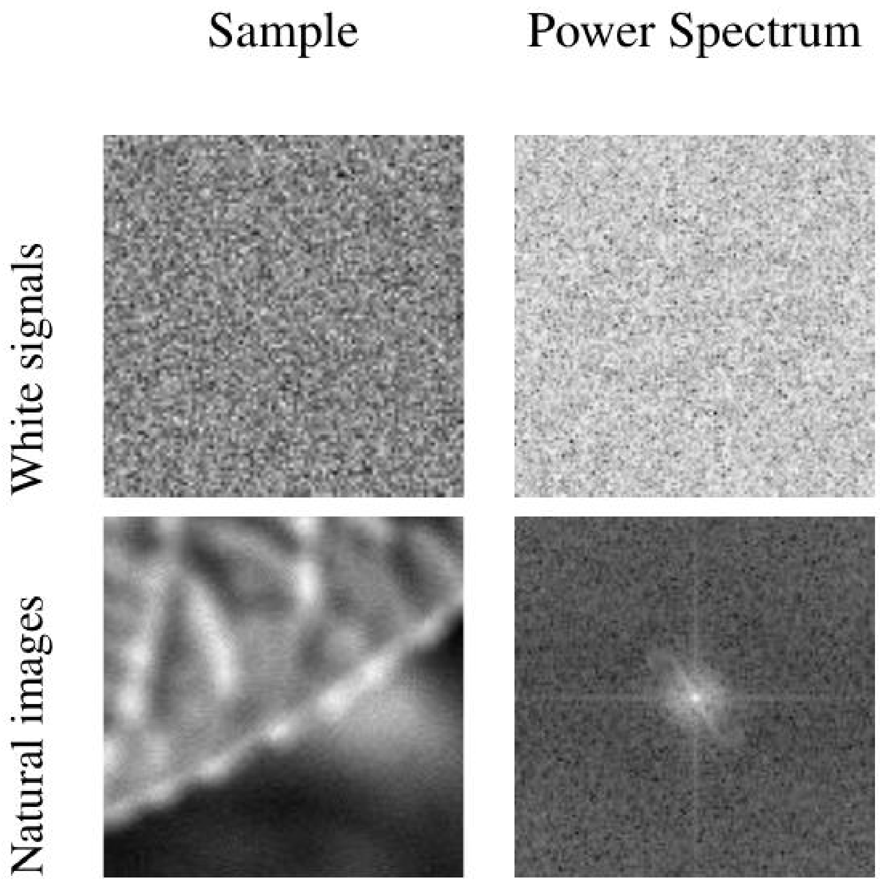
$$E[PS(X)) = F(\mu^2[X])$$

Contents

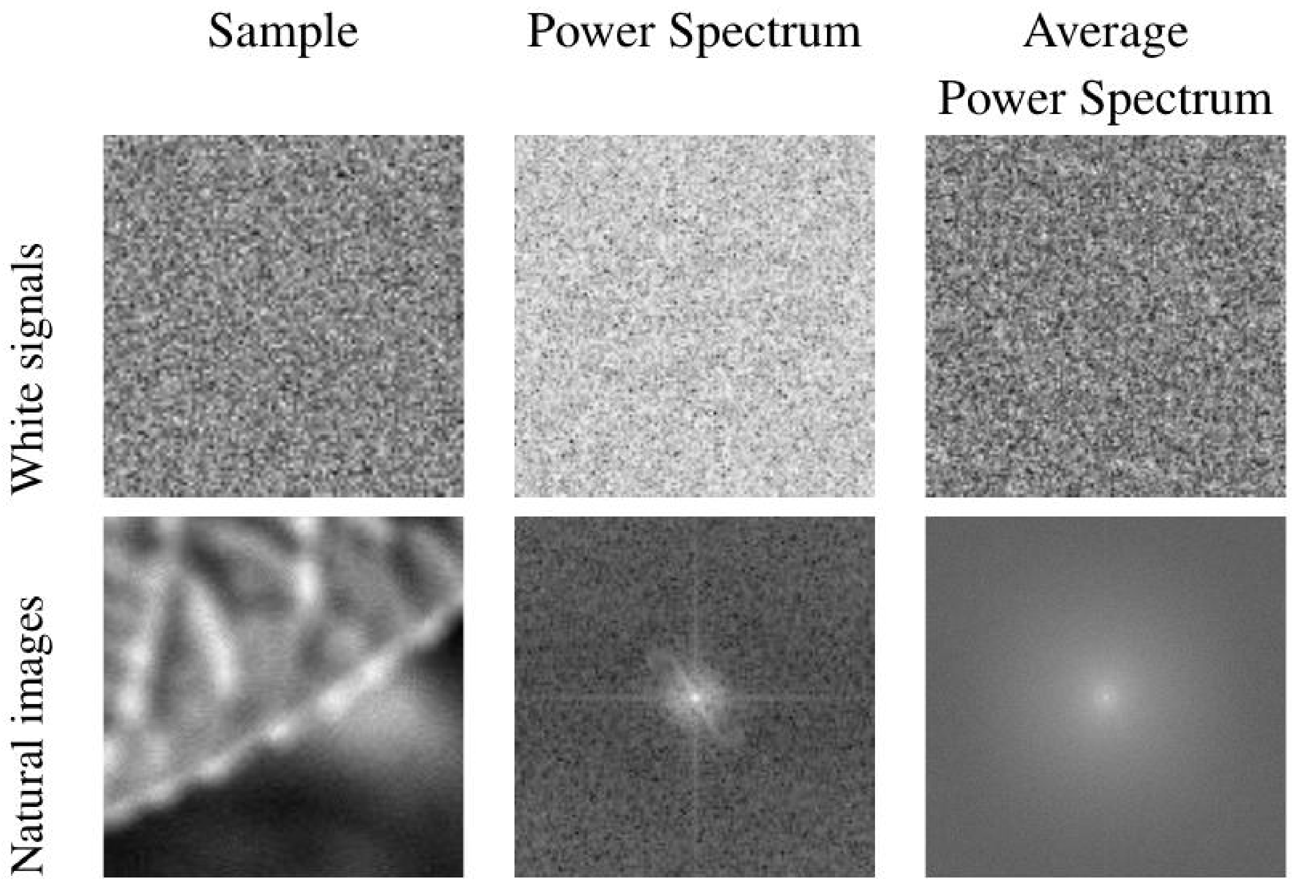
Power Spectrum of Natural Images



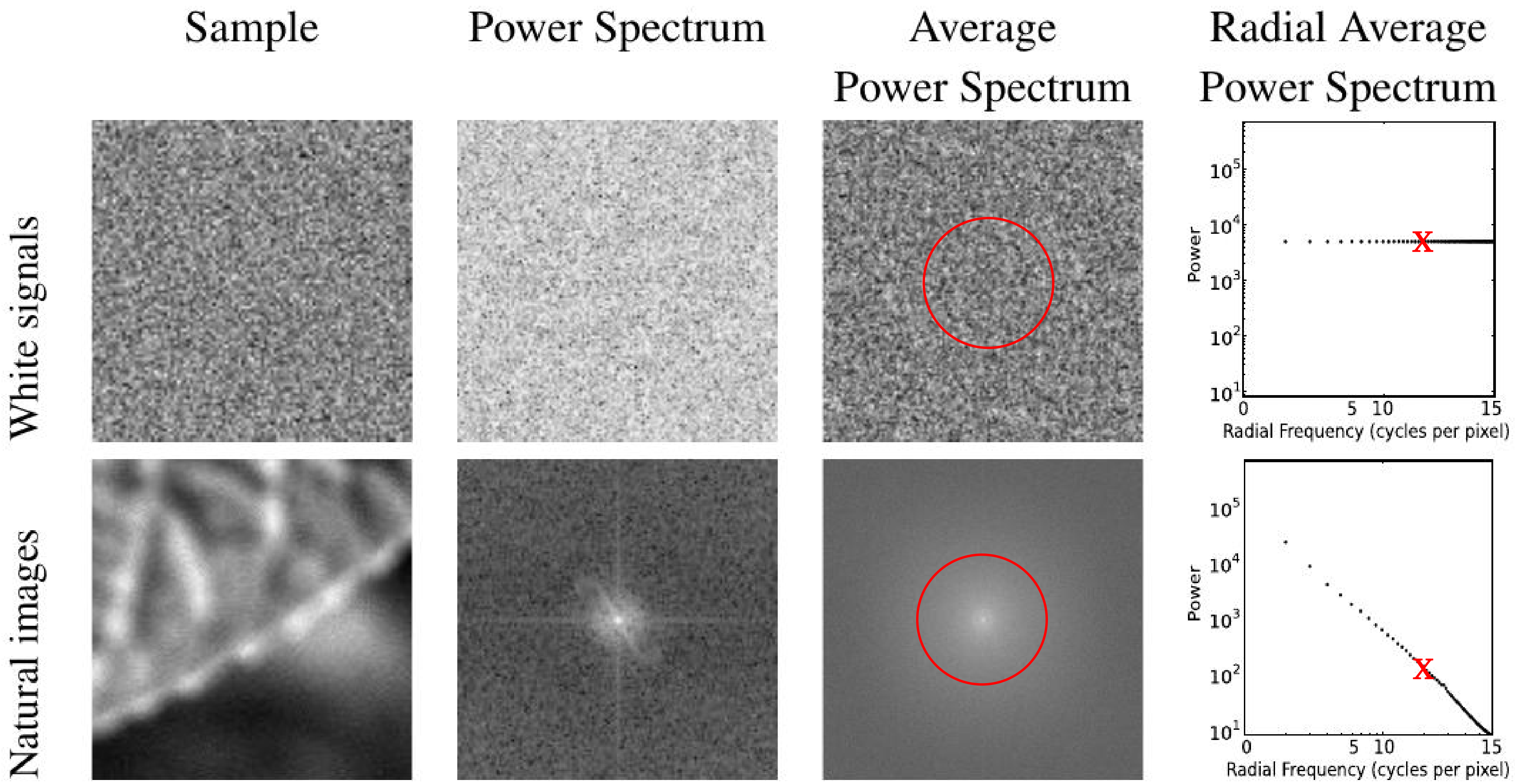
Power Spectrum of Natural Images



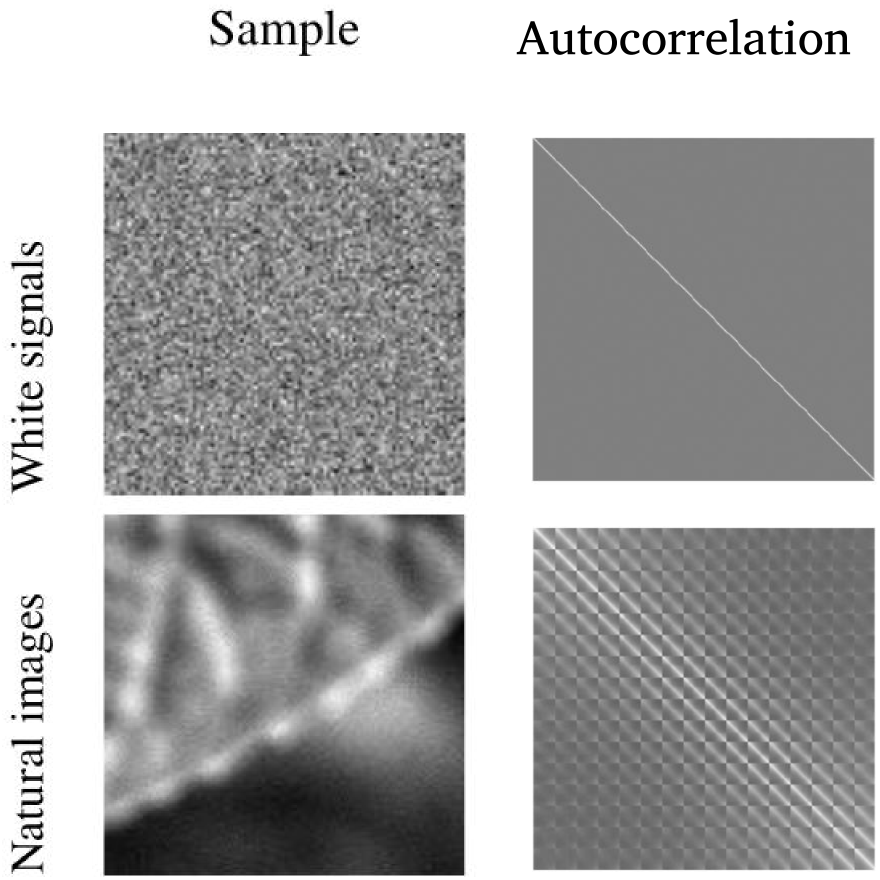
Power Spectrum of Natural Images



Power Spectrum of Natural Images



Auto correlation of natural images

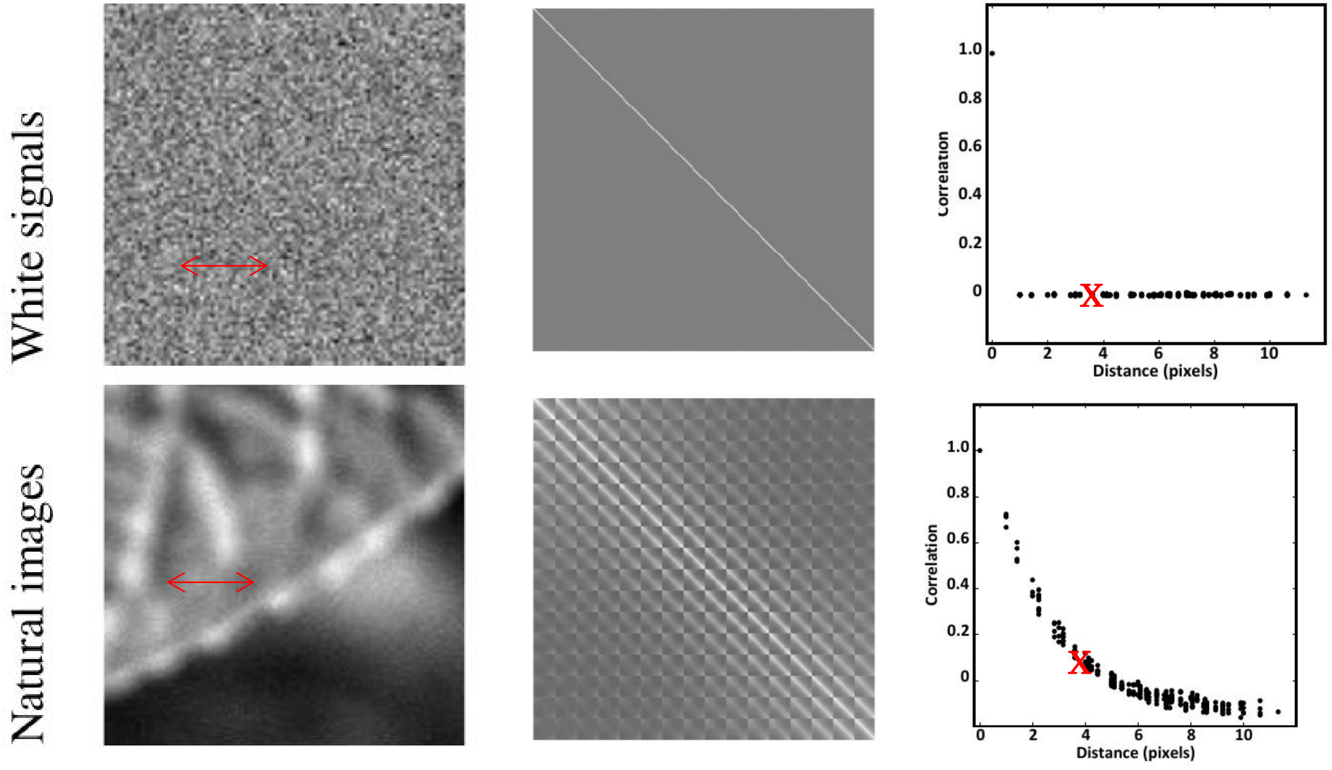


Auto correlation of natural images

Sample

Autocorrelation

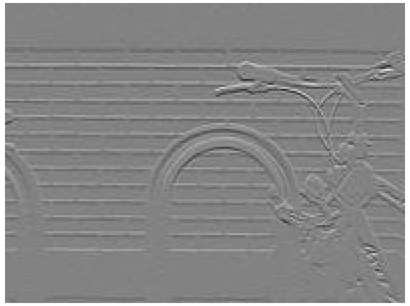
Autocorrelation in function of pixel distance



Sobel Filters: Introduction

- It performs a 2-D spatial gradient measurement on images in order to emphasizes edges.
- Pairs of convolution kernels (K_x and K_y) designed to respond maximally to edges running vertically and horizontally relative to the pixel grid

$G_x = I * K_x$ $\qquad$ $G_y = I * K_y$



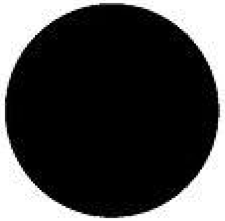
Sobel Filters: Introduction

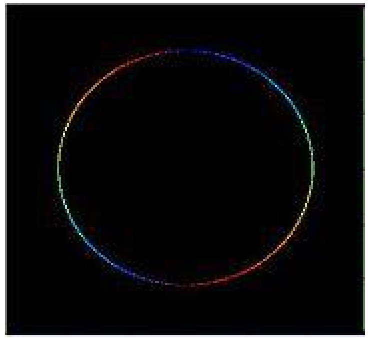
- Magnitude of the gradient:

$$|G| = \sqrt{(G_x^2 + G_y^2)}$$

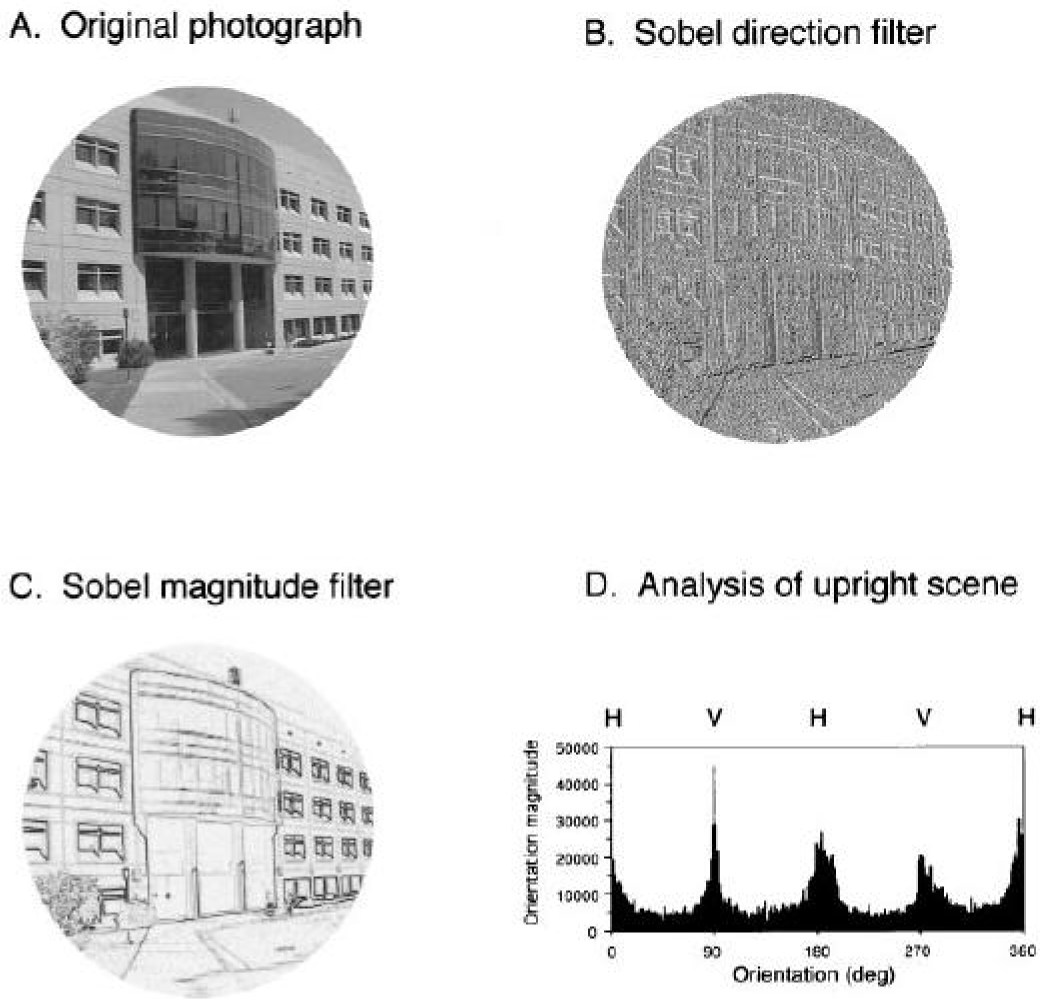
- Angle of the gradient:

$$\theta = \arctan(G_y/G_x)$$



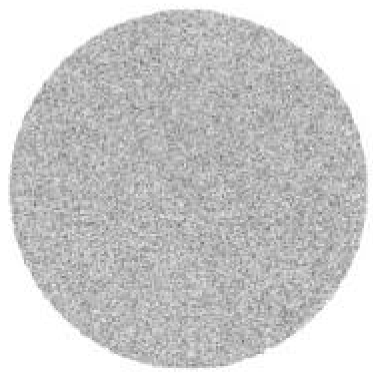


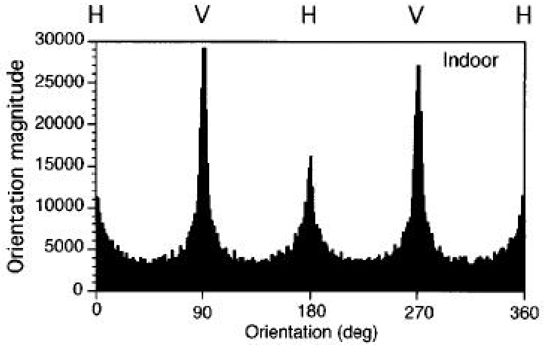
Analysis of edges orientations

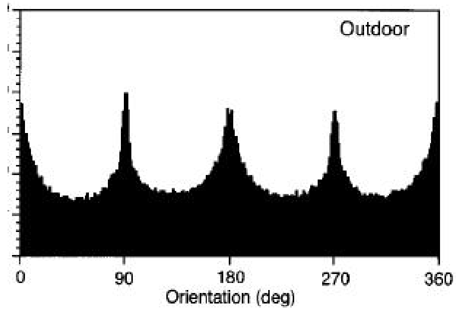


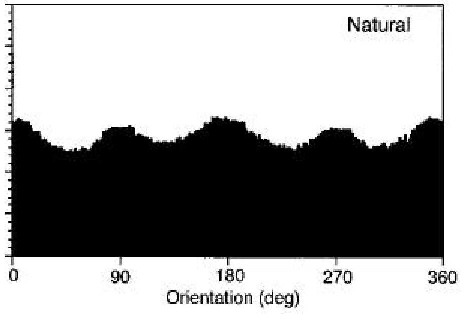
The distribution of oriented contoursin the real world, Coppola et al

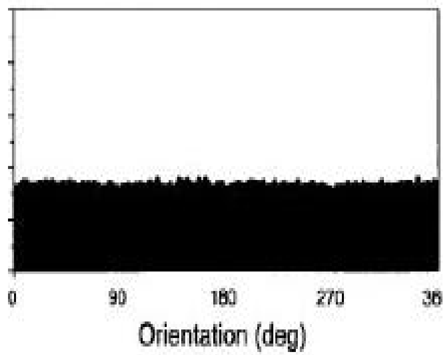
Analysis of edges orientations

Indoor	Outdoor	Natural	White Noise
			





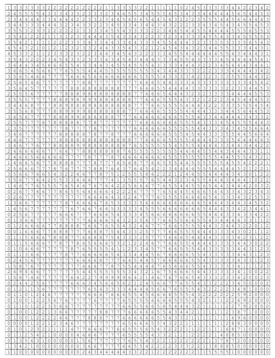




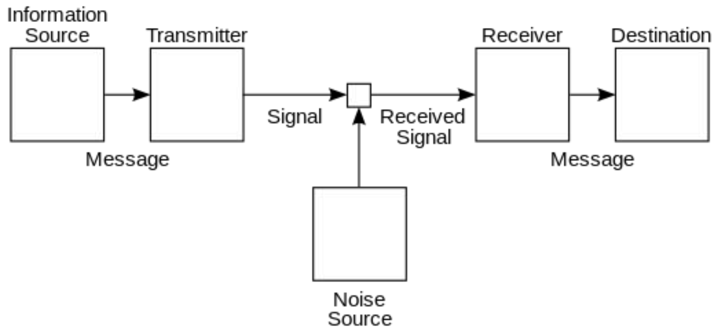
The distribution of oriented contoursin the real world, Coppola et al

Summary

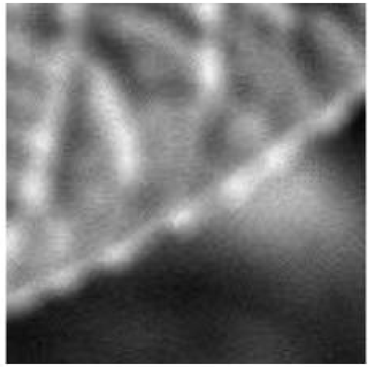
1. What is vision?

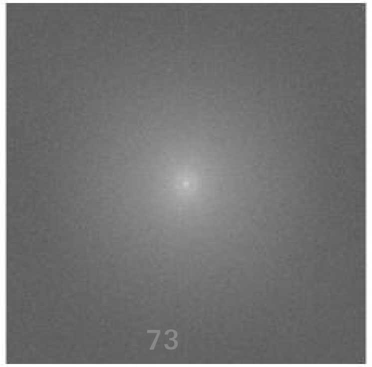


2. The redundancy reduction hypothesis



3. Natural images statistics





Bibliography

- Marr, David, Vision: A computational investigation into the human representation and processing of visual information, 1982
- Shannon, A mathematical theory of communication, 1948
- Marc URO, Basic Concepts in information theory, http://www-public.tem-tsp.eu/~uro/cours-pdf/poly.pdf
- Barlow HB Possible principles underlying the transformation of sensory messages, 1961
- Attneave, Some informational aspects of visual perception, 1954

Bibliography (cont)

- Hyvarinen et al, Natural Image Statistics, 2009
- Kersten Predictability and redundancy of natural images, 1987
- Attneave, Some informational aspects of visual perception, 1954
- Attick, Could information theory provide an ecological theory of sensory processing?, 1992
- Ruderman, The statistics of natural images, 1994
- Coppola et all, The distribution of oriented contoursin the real world, 2001

Extra Slides