Natural Images

Giovanni Marelli

Göttingen 27.01.10

Infomation in Natural Images

Infomation in Natural Images

- How much information do they carry?

Infomation in Natural Images

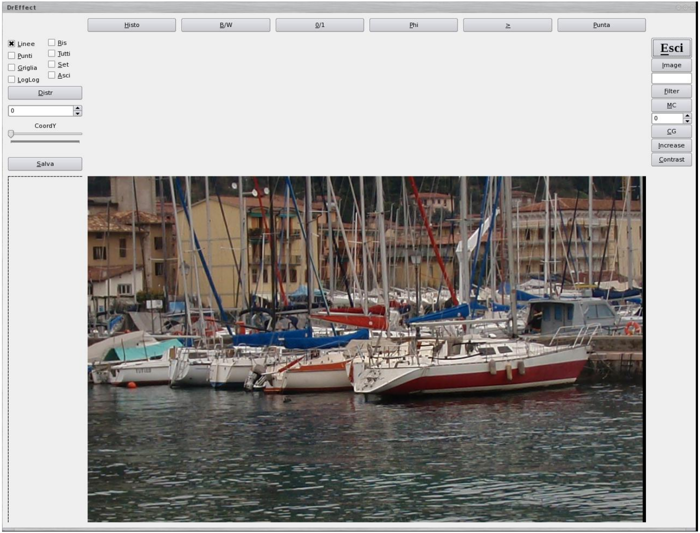
- How much information do they carry?
- Are they organized in some specific structures?

Infomation in Natural Images

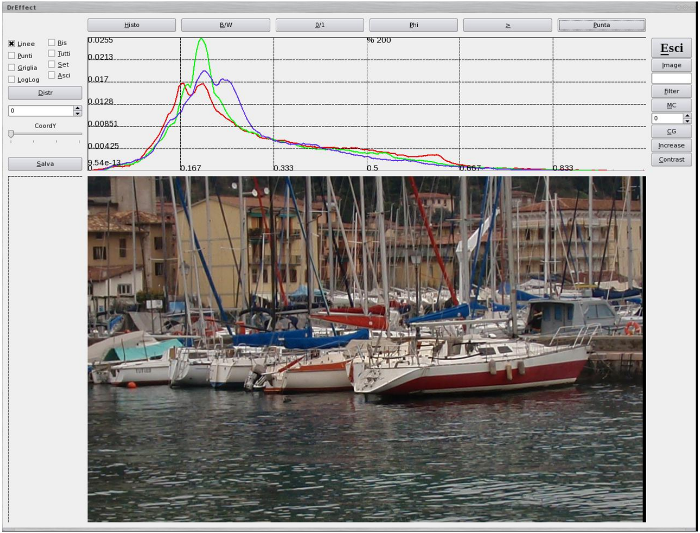
- How much information do they carry?
- Are they organized in some specific structures?
- Do they belong to a specific class of universality?

A picture is composed by $n \times m$ matrix of pixels

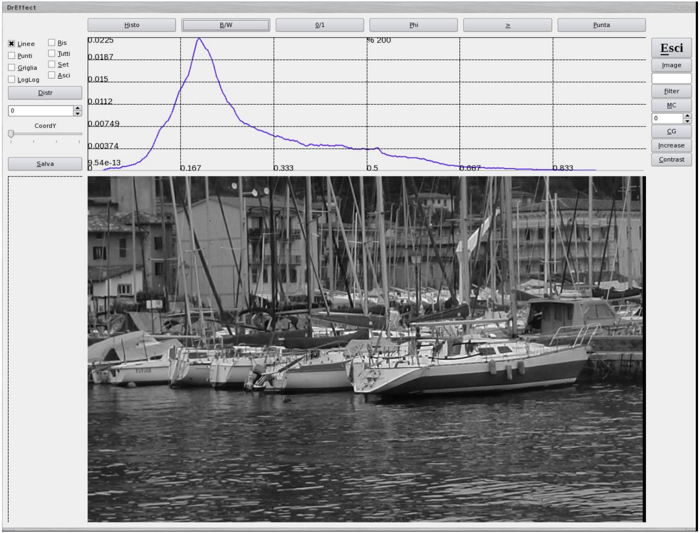
Every pixel contains an array of R, G, B, (eventually α as transparency) which have values within $0, 255$ which define the intensity I of the picture.



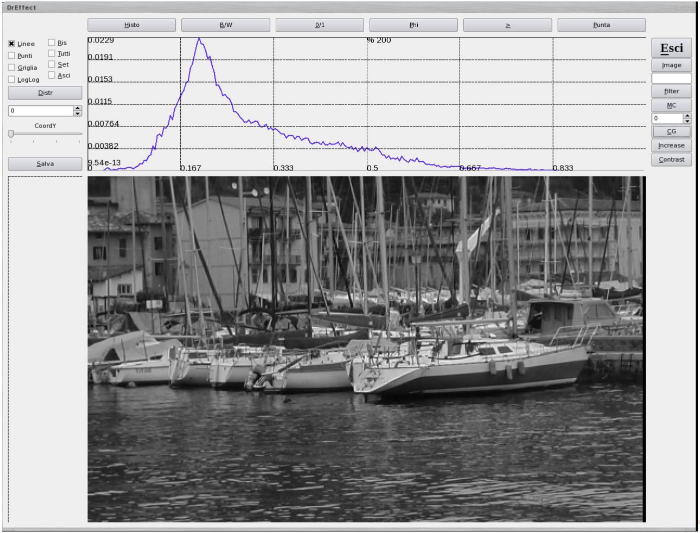
The hystogram shows the distribution of the intensities in the picture



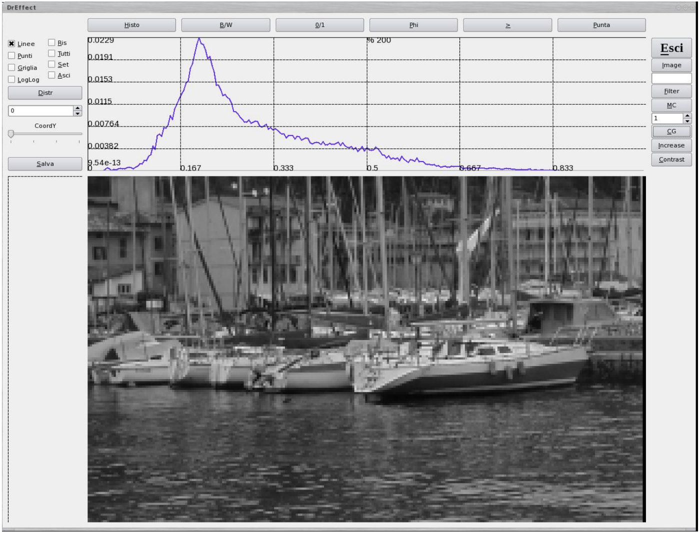
We turn the image into a gray scale image



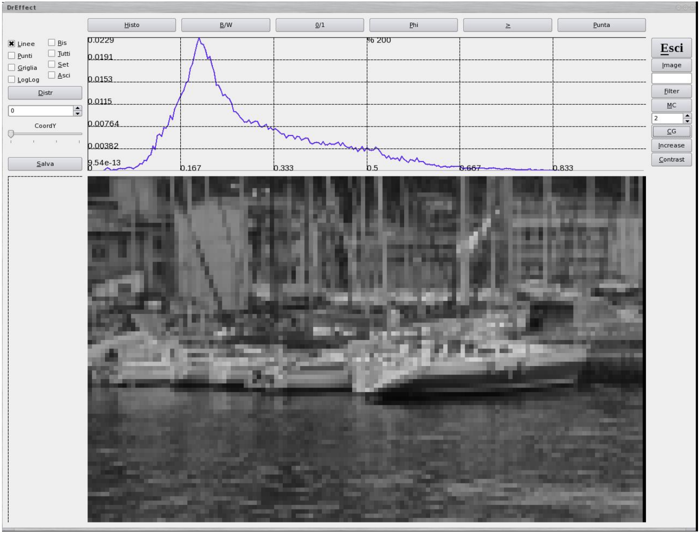
We turn the image into a gray scale image for different coarse grained pixel: 1:4



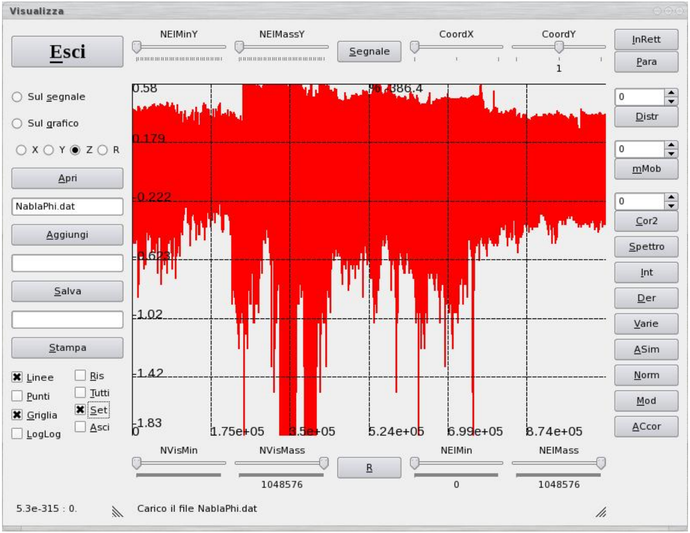
We turn the image into a gray scale image for different coarse grained pixel: 1:16



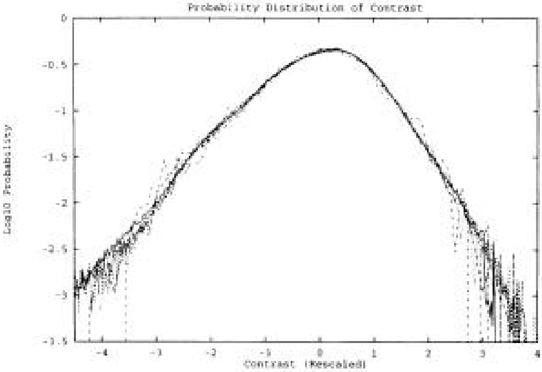
We turn the image into a gray scale image for different coarse grained pixel: 1:32



We define the contrast as $\phi() := \log(I(\boldsymbol{x})/I_0)$ where $I_0 : N(\phi < 0) = N(\phi > 0)$

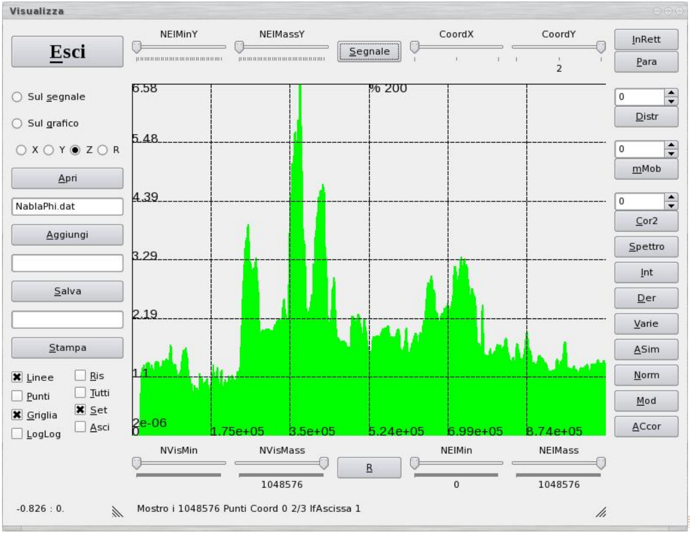


If we calculate the probability distribution
we see that for different coarse grained pixel the plot shows long tails



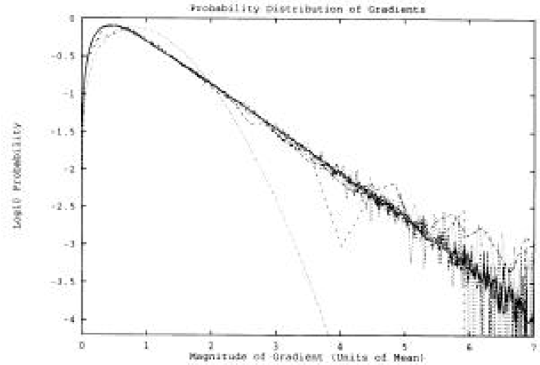
If we calculate ist gradient $|\nabla\phi|$

The distribution is quite precisely exponential

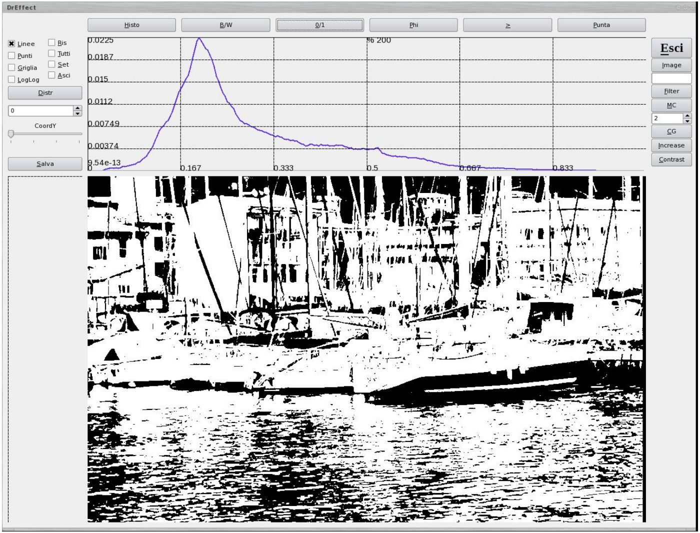


If we calculate ist gradient $|\nabla\phi|$

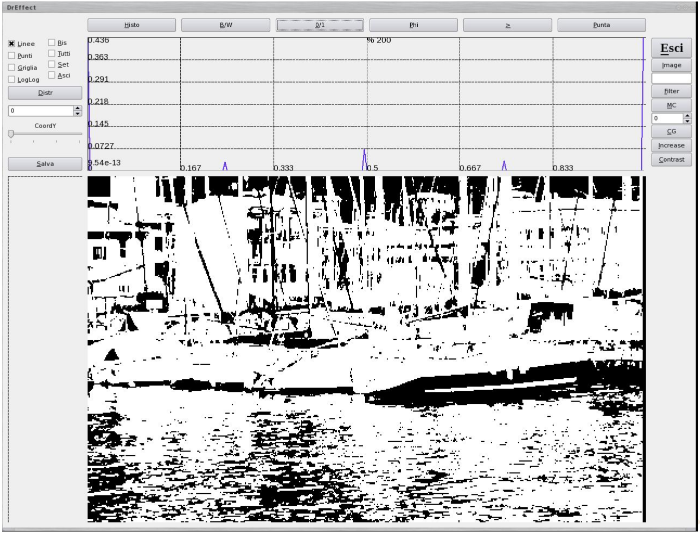
The distribution is quite precisely exponential



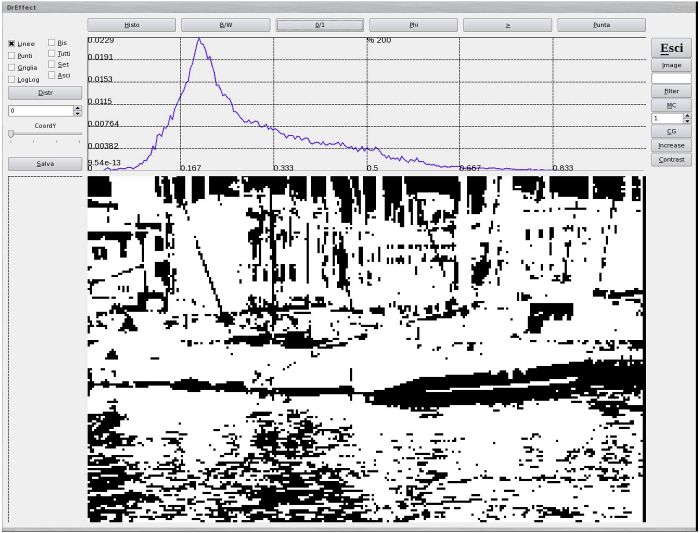
We calculate the median of the distribution and assign 0 (black) for every intensity below the median and 1 (white) for every intesities above. The pixel are equally populated



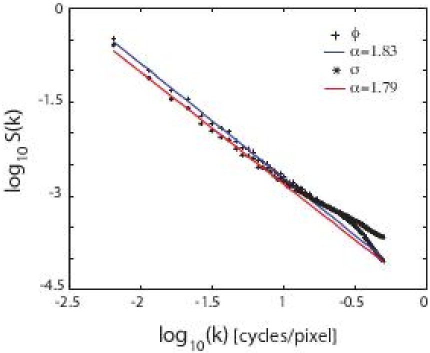
We calculate the median of the distribution and assign 0 (black) for every intensity below the median and 1 (white) for every intesities above. The pixel are equally populated Coarse grained: 1:4



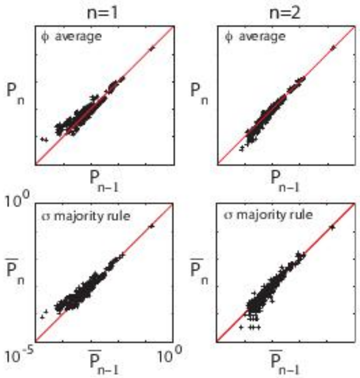
We calculate the median of the distribution and assign 0 (black) for every intensity below the median and 1 (white) for every intesities above. The pixel are equally populated Coarse grained: 1:16



The power spectrum in the loglog plot shows a clear power law



The full distribution is invariant to block scaling if we coarse-grain and quantize (top) or quantize and coarse-grain (bottom)



We consider a receptor cell with a contrast noise of variance σ^2 and a Nyquist frequency of k_c.

The power spectrum is connected with the autocorrelation (Th. Wiener-Khintchin)
$< \phi(\boldsymbol{x})\phi(\boldsymbol{x}') > := \int \frac{\mathrm{d}^2 k}{(2\pi)^2} S_\phi(\boldsymbol{k}) e^{\imath \boldsymbol{k}(\boldsymbol{x}-\boldsymbol{x}')}$

Among all the distributions, the gaussian has the maximum entropy. In this assumption one can connect the autocorrelation of the picture with the information conveyed
$I \leq \frac{\pi N \pi/2}{k_c^2} \int \frac{\mathrm{d}^2 k}{(2\pi)^2} \log_2 \left(1 + \frac{(1-k/k_c)^2}{\pi^2 \sigma^2} k_c^2 S_\phi(k)\right)$

The signal to noise ration of a single cell is
$R_{SNR} = \frac{1}{\sigma^2} \int \frac{\mathrm{d}^2 k}{(2\pi)^2} (1 - k/k_c)^1 S_\phi(k)$
$I \leq \frac{1}{2} N \frac{\pi}{2} \int_0^1 \mathrm{d}x x \log_2 \left(1 + \frac{\eta(\eta+1)(\eta+2)}{\pi} R_{SNR} \frac{(10x)^2}{x^{2-\eta}}\right)$

A single receptor can produce a spike when the signal overcome a certain threshold
The upper bound we have shown tells that the information contained in a natural image is less than a bit per receptor.

A channel that transmit the information to the brain can have capacity lower than 1/2 bit/receptor to convey all the information

D.L. Ruderman, dissertation, University of California at Berkley, 1993

We consider a receptor cell with a contrast noise of variance σ^2 and a Nyquist frequency of k_c.

The power spectrum is connected with the autocorrelation (Th. Wiener-Khintchin)
$< \phi(\boldsymbol{x})\phi(\boldsymbol{x}') > := \int \frac{\mathrm{d}^2 k}{(2\pi)^2} S_\phi(\boldsymbol{k}) e^{\imath \boldsymbol{k}(\boldsymbol{x}-\boldsymbol{x}')}$

Among all the distributions, the gaussian has the maximum entropy. In this assumption one can connect the autocorrelation of the picture with the information conveyed
$I \leq \frac{\pi N \pi/2}{k_c^2} \int \frac{\mathrm{d}^2 k}{(2\pi)^2} \log_2 \left(1 + \frac{(1-k/k_c)^2}{\pi^2 \sigma^2} k_c^2 S_\phi(k)\right)$

The signal to noise ration of a single cell is
$R_{SNR} = \frac{1}{\sigma^2} \int \frac{\mathrm{d}^2 k}{(2\pi)^2} (1 - k/k_c)^1 S_\phi(k)$
$I \leq \frac{1}{2} N \frac{\pi}{2} \int_0^1 \mathrm{d}x x \log_2 \left(1 + \frac{\eta(\eta+1)(\eta+2)}{\pi} R_{SNR} \frac{(10x)^2}{x^{2-\eta}}\right)$

A single receptor can produce a spike when the signal overcome a certain threshold
The upper bound we have shown tells that the information contained in a natural image is less than a bit per receptor.

A channel that transmit the information to the brain can have capacity lower than 1/2 bit/receptor to convey all the information

D.L. Ruderman, dissertation, University of California at Berkley, 1993

We consider a receptor cell with a contrast noise of variance σ^2 and a Nyquist frequency of k_c.

The power spectrum is connected with the autocorrelation (Th. Wiener-Khintchin)
$< \phi(\boldsymbol{x})\phi(\boldsymbol{x}') >:= \int \frac{\mathrm{d}^2 k}{(2\pi)^2} S_\phi(\boldsymbol{k}) e^{\imath \boldsymbol{k}(\boldsymbol{x}-\boldsymbol{x}')}$

Among all the distributions, the gaussian has the maximum entropy. In this assumption one can connect the autocorrelation of the picture with the information conveyed
$I \leq \frac{\pi N \pi/2}{k_c^2} \int \frac{\mathrm{d}^2 k}{(2\pi)^2} \log_2 \left(1 + \frac{(1-k/k_c)^2}{\pi^2 \sigma^2} k_c^2 S_\phi(k)\right)$

The signal to noise ration of a single cell is
$R_{SNR} = \frac{1}{\sigma^2} \int \frac{\mathrm{d}^2 k}{(2\pi)^2} (1 - k/k_c)^1 S_\phi(k)$
$I \leq \frac{1}{2} N \frac{\pi}{2} \int_0^1 \mathrm{d}x x \log_2 \left(1 + \frac{\eta(\eta+1)(\eta+2)}{\pi} R_{SNR} \frac{(10x)^2}{x^{2-\eta}}\right)$

A single receptor can produce a spike when the signal overcome a certain threshold
The upper bound we have shown tells that the information contained in a natural image is less than a bit per receptor.

A channel that transmit the information to the brain can have capacity lower than 1/2 bit/receptor to convey all the information

D.L. Ruderman, dissertation, University of California at Berkley, 1993

We consider a receptor cell with a contrast noise of variance σ^2 and a Nyquist frequency of k_c.

The power spectrum is connected with the autocorrelation (Th. Wiener-Khintchin)
$< \phi(\boldsymbol{x})\phi(\boldsymbol{x}') >:= \int \frac{\mathrm{d}^2 k}{(2\pi)^2} S_\phi(\boldsymbol{k}) e^{\imath \boldsymbol{k}(\boldsymbol{x}-\boldsymbol{x}')}$

Among all the distributions, the gaussian has the maximum entropy. In this assumption one can connect the autocorrelation of the picture with the information conveyed
$I \leq \frac{\pi N \pi/2}{k_c^2} \int \frac{\mathrm{d}^2 k}{(2\pi)^2} \log_2 \left(1 + \frac{(1-k/k_c)^2}{\pi^2 \sigma^2} k_c^2 S_\phi(k) \right)$

The signal to noise ration of a single cell is
$R_{SNR} = \frac{1}{\sigma^2} \int \frac{\mathrm{d}^2 k}{(2\pi)^2} (1-k/k_c)^1 S_\phi(k)$
$I \leq \frac{1}{2} N \frac{\pi}{2} \int_0^1 \mathrm{d}x x \log_2 \left(1 + \frac{\eta(\eta+1)(\eta+2)}{\pi} R_{SNR} \frac{(10x)^2}{x^{2-\eta}} \right)$

A single receptor can produce a spike when the signal overcome a certain threshold
The upper bound we have shown tells that the information contained in a natural image is less than a bit per receptor.

A channel that transmit the information to the brain can have capacity lower than 1/2 bit/receptor to convey all the information

D.L. Ruderman, dissertation, University of California at Berkley, 1993

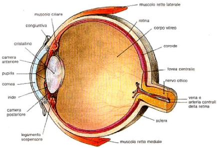
We consider a receptor cell with a contrast noise of variance σ^2 and a Nyquist frequency of k_c.

The power spectrum is connected with the autocorrelation (Th. Wiener-Khintchin)
$< \phi(\boldsymbol{x})\phi(\boldsymbol{x}') > := \int \frac{\mathrm{d}^2 k}{(2\pi)^2} S_\phi(\boldsymbol{k}) e^{\imath \boldsymbol{k}(\boldsymbol{x}-\boldsymbol{x}')}$

Among all the distributions, the gaussian has the maximum entropy. In this assumption one can connect the autocorrelation of the picture with the information conveyed
$I \leq \frac{\pi N \pi/2}{k_c^2} \int \frac{\mathrm{d}^2 k}{(2\pi)^2} \log_2 \left(1 + \frac{(1-k/k_c)^2}{\pi^2\sigma^2} k_c^2 S_\phi(k)\right)$

The signal to noise ration of a single cell is
$R_{SNR} = \frac{1}{\sigma^2} \int \frac{\mathrm{d}^2 k}{(2\pi)^2} (1 - k/k_c)^1 S_\phi(k)$
$I \leq \frac{1}{2} N \frac{\pi}{2} \int_0^1 \mathrm{d}x x \log_2 \left(1 + \frac{\eta(\eta+1)(\eta+2)}{\pi} R_{SNR} \frac{(10x)^2}{x^{2-\eta}}\right)$

A single receptor can produce a spike when the signal overcome a certain threshold
The upper bound we have shown tells that the information contained in a natural image is less than a bit per receptor.



A channel that transmit the information to the brain can have capacity lower than 1/2 bit/receptor to convey all the information

D.L. Ruderman, dissertation, University of California at Berkley, 1993

Do the black/white pixels remember to the $(\uparrow, \downarrow)$ of the Ising model?

A region of L square pixel has $2^{L\times L}$ states, considering the entropy:
$S(3\times 3) = 6.580 \pm 0.003 < 9\text{bit}$,
$S(4\times 4) = 11.154 \pm 0.002\text{bit}$

$$Z(T) = \frac{1}{\Delta}\int \mathrm{d}E e^{S(E)-E/T}$$

To the sample s of $L\times L$ we associate an energy. The Boltzmann probability distribution for a given "Temperature" is

$$P(s) = \frac{1}{Z(T)} e^{-E(s)/T}$$

$$S(T) = -\sum_s P(s) \log P(s)$$

One can define new different quantities, the heat capacity

$$C(T) = T\partial_T S(T) = \frac{<(\delta E(s))^2>_T}{T^2}$$

A natural image will be compared to a Monte Carlo simulation of the same system size of a Ising system

Do the black/white pixels remember to the $(\uparrow, \downarrow)$ of the Ising model?

A region of L square pixel has $2^{L\times L}$ states, considering the entropy:
$S(3\times 3) = 6.580 \pm 0.003 < 9\mathrm{bit}$,
$S(4\times 4) = 11.154 \pm 0.002\mathrm{bit}$

$$Z(T) = \frac{1}{\Delta}\int \mathrm{d}E e^{S(E)-E/T}$$

To the sample s of $L\times L$ we associate an energy. The Boltzmann probability distribution for a given "Temperature" is

$$P(s) = \frac{1}{Z(T)} e^{-E(s)/T}$$

$$S(T) = -\sum_s P(s)\log P(s)$$

One can define new different quantities, the heat capacity

$$C(T) = T\partial_T S(T) = \frac{<(\delta E(s))^2>_T}{T^2}$$

A natural image will be compared to a Monte Carlo simulation of the same system size of a Ising system

Do the black/white pixels remember to the $(\uparrow, \downarrow)$ of the Ising model?

A region of L square pixel has $2^{L\times L}$ states, considering the entropy:
$S(3\times 3) = 6.580 \pm 0.003 < 9\text{bit}$,
$S(4\times 4) = 11.154 \pm 0.002\text{bit}$

$$Z(T) = \frac{1}{\Delta}\int \mathrm{d}E e^{S(E)-E/T}$$

To the sample s of $L\times L$ we associate an energy. The Boltzmann probability distribution for a given "Temperature" is

$$P(s) = \frac{1}{Z(T)} e^{-E(s)/T}$$

$$S(T) = -\sum_s P(s)\log P(s)$$

One can define new different quantities, the heat capacity

$$C(T) = T\partial_T S(T) = \frac{<(\delta E(s))^2>_T}{T^2}$$

A natural image will be compared to a Monte Carlo simulation of the same system size of a Ising system

Do the black/white pixels remember to the $(\uparrow, \downarrow)$ of the Ising model?

A region of L square pixel has $2^{L\times L}$ states, considering the entropy:
$S(3\times 3) = 6.580 \pm 0.003 < 9\text{bit}$,
$S(4\times 4) = 11.154 \pm 0.002\text{bit}$

$$Z(T) = \frac{1}{\Delta}\int \mathrm{d}E e^{S(E)-E/T}$$

To the sample s of $L\times L$ we associate an energy. The Boltzmann probability distribution for a given "Temperature" is

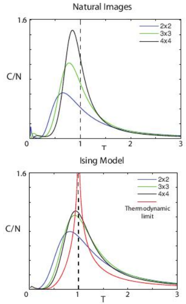
$$P(s) = \frac{1}{Z(T)} e^{-E(s)/T}$$

$$S(T) = -\sum_s P(s)\log P(s)$$

One can define new different quantities, the heat capacity

$$C(T) = T\partial_T S(T) = \frac{<(\delta E(s))^2>_T}{T^2}$$

A natural image will be compared to a Monte Carlo simulation of the same system size of a Ising system



A patch in a picture is as a word in a book

Zipf's law
$P_r \propto 1/r^\alpha$

$\alpha \simeq 1$, r rank

$p_r = A/r^\alpha$

If we identify the energy for a given rank of a Boltzmann distribution the energy for the rank r is $E_r = \alpha \ln r = \ln AZ$

In the thermodynamic limit the density of states $\rho(E) \simeq |\mathrm{d}E/\mathrm{d}r|^{-1} = r/\alpha$ which gives
$\rho(E) = \frac{1}{\alpha}(AZ)^{1/\alpha} e^{E/\alpha} e^{E/\alpha}$ $\quad S_{Zipf}(E) = E/\alpha + \mathrm{const}$

A patch in a picture is as a word in a book

Zipf's law
$P_r \propto 1/r^\alpha$

$\alpha \simeq 1$, r rank

$p_r = A/r^\alpha$

If we identify the energy for a given rank of a Boltzmann distribution the energy for the rank r is $E_r = \alpha \ln r = \ln AZ$

In the thermodynamic limit the density of states
$\rho(E) \simeq |\mathrm{d}E/\mathrm{d}r|^{-1} = r/\alpha$ which gives
$\rho(E) = \frac{1}{\alpha}(AZ)^{1/\alpha} e^{E/\alpha} e^{E/\alpha}$ $\quad S_{Zipf}(E) = E/\alpha + \mathrm{const}$

A patch in a picture is as a word in a book

Zipf's law
$P_r \propto 1/r^\alpha$

$\alpha \simeq 1$, r rank

$p_r = A/r^\alpha$

If we identify the energy for a given rank of a Boltzmann distribution the energy for the rank r is $E_r = \alpha \ln r = \ln AZ$

In the thermodynamic limit the density of states
$\rho(E) \simeq |\mathrm{d}E/\mathrm{d}r|^{-1} = r/\alpha$ which gives
$\rho(E) = \frac{1}{\alpha}(AZ)^{1/\alpha} e^{E/\alpha} e^{E/\alpha}$ $\quad S_{Zipf}(E) = E/\alpha + \mathrm{const}$

A patch in a picture is as a word in a book

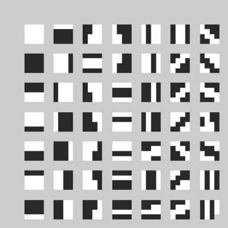
Zipf's law
$P_r \propto 1/r^\alpha$

$\alpha \simeq 1$, r rank

$p_r = A/r^\alpha$

If we identify the energy for a given rank of a Boltzmann distribution the energy for the rank r is $E_r = \alpha \ln r = \ln AZ$

In the thermodynamic limit the density of states
$\rho(E) \simeq |\mathrm{d}E/\mathrm{d}r|^{-1} = r/\alpha$ which gives
$\rho(E) = \frac{1}{\alpha}(AZ)^{1/\alpha} e^{E/\alpha} e^{E/\alpha}$ $\quad S_{Zipf}(E) = E/\alpha + \mathrm{const}$



Most common patches

The natural images repeat specific patterns and contain more information.

Can one increase the resolution? Cheating in physics

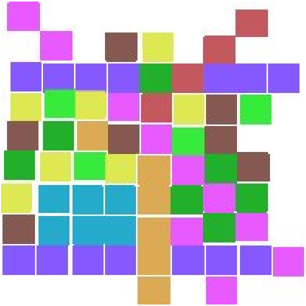
The previous analysis shows a long correlation between the structure

Considering many natural images one can, as was done to compress the natual languages, finding the two dimensional correlation between the words.

The correlation should be, as we already said, both universal and dependent on the photographer and on the landscape.
Once we have calculated the correlation function for our picture, can we substitute different coarsed grained order of pathces?
Even the Fourier space would be sensitive to repeated structures (Bragg's planes)

The natural images repeat specific patterns and contain more information.

Can one increase the resolution? Cheating in physics



The previous analysis shows a long correlation between the structure

Considering many natural images one can, as was done to compress the natual languages, finding the two dimensional correlation between the words.

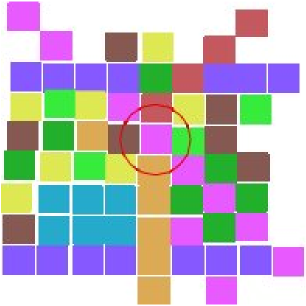
The correlation should be, as we already said, both universal and dependent on the photographer and on the landscape.

Once we have calculated the correlation function for our picture, can we substitute different coarsed grained order of pathces?

Even the Fourier space would be sensitive to repeated structures (Bragg's planes)

The natural images repeat specific patterns and contain more information.

Can one increase the resolution? Cheating in physics



The previous analysis shows a long correlation between the structure

Considering many natural images one can, as was done to compress the natual languages, finding the two dimensional correlation between the words.

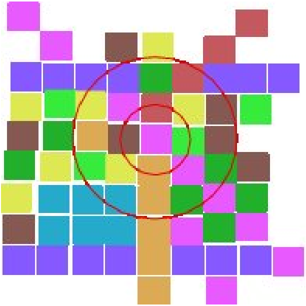
The correlation should be, as we already said, both universal and dependent on the photographer and on the landscape.

Once we have calculated the correlation function for our picture, can we substitute different coarsed grained order of pathces?

Even the Fourier space would be sensitive to repeated structures (Bragg's planes)

The natural images repeat specific patterns and contain more information.

Can one increase the resolution? Cheating in physics



The previous analysis shows a long correlation between the structure

Considering many natural images one can, as was done to compress the natual languages, finding the two dimensional correlation between the words.

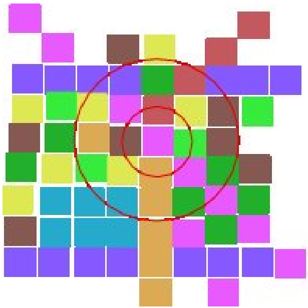
The correlation should be, as we already said, both universal and dependent on the photographer and on the landscape.

Once we have calculated the correlation function for our picture, can we substitute different coarsed grained order of pathces?

Even the Fourier space would be sensitive to repeated structures (Bragg's planes)

The natural images repeat specific patterns and contain more information.

Can one increase the resolution? Cheating in physics



The previous analysis shows a long correlation between the structure

Considering many natural images one can, as was done to compress the natual languages, finding the two dimensional correlation between the words.

The correlation should be, as we already said, both universal and dependent on the photographer and on the landscape.

Once we have calculated the correlation function for our picture, can we substitute different coarsed grained order of pathces?

Even the Fourier space would be sensitive to repeated structures (Bragg's planes)

Proving an increment in resolution

To test different methods to increase the resolution one could

Resize the original image to a coarse-grained one

Try different methods to create one with more resolution

Compress the original images into its more probable pathces

Compare the created images with the original one

$C_c = \frac{1}{N^4} \sum_{ijkl} (p_{ij} - p_{kl})^2$,
$C_a = \frac{1}{N^2} \sum_{ij} (p_{ij} - p_{ij})^2$

Will be the information increased whithin the error?

Proving an increment in resolution

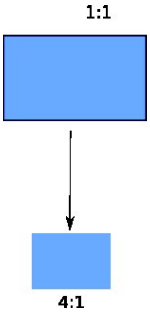
To test different methods to increase the resolution one could
Resize the original image to a coarse-grained one

Try different methods to create one with more resolution

Compress the original images into its more probable pathces

Compare the created images with the original one

$C_c = \frac{1}{N^4} \sum_{ijkl} (p_{ij} - p_{kl})^2$,
$C_a = \frac{1}{N^2} \sum_{ij} (p_{ij} - p_{ij})^2$



Will be the information increased whithin the error?

Proving an increment in resolution

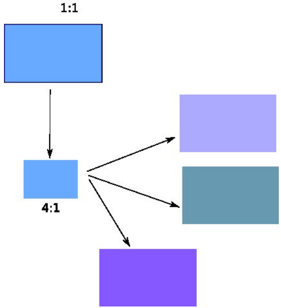
To test different methods to increase the resolution one could Resize the original image to a coarse-grained one

Try different methods to create one with more resolution

Compress the original images into its more probable pathces

Compare the created images with the original one

$C_c = \frac{1}{N^4} \sum_{ijkl} (p_{ij} - p_{kl})^2$,
$C_a = \frac{1}{N^2} \sum_{ij} (p_{ij} - p_{ij})^2$



Will be the information increased whithin the error?

Proving an increment in resolution

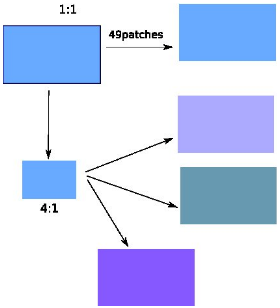
To test different methods to increase the resolution one could Resize the original image to a coarse-grained one

Try different methods to create one with more resolution

Compress the original images into its more probable pathces

Compare the created images with the original one

$C_c = \frac{1}{N^4} \sum_{ijkl} (p_{ij} - p_{kl})^2,$
$C_a = \frac{1}{N^2} \sum_{ij} (p_{ij} - p_{ij})^2$



Will be the information increased whithin the error?

Proving an increment in resolution

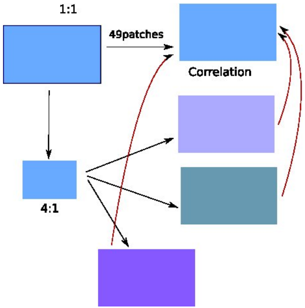
To test different methods to increase the resolution one could Resize the original image to a coarse-grained one

Try different methods to create one with more resolution

Compress the original images into its more probable pathces

Compare the created images with the original one

$C_c = \frac{1}{N^4} \sum_{ijkl} (p_{ij} - p_{kl})^2,$
$C_a = \frac{1}{N^2} \sum_{ij} (p_{ij} - p_{ij})^2$



Will be the information increased whithin the error?

Proving an increment in resolution

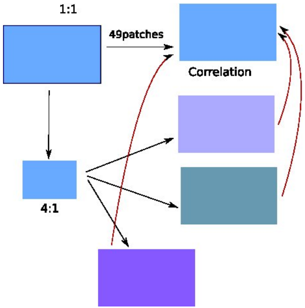
To test different methods to increase the resolution one could Resize the original image to a coarse-grained one

Try different methods to create one with more resolution

Compress the original images into its more probable pathces

Compare the created images with the original one

$C_c = \frac{1}{N^4} \sum_{ijkl} (p_{ij} - p_{kl})^2,$
$C_a = \frac{1}{N^2} \sum_{ij} (p_{ij} - p_{ij})^2$



Will be the information increased whithin the error?

Thank you for your attention

Aknowledgments
Wolfgang Keil (tutor)

References

Thermodynamics of natural images

Stephens, Greg J; Mora, Thierry; Tkacik, Gasper; Bialek, William

http://arxiv.org/abs/0806.2694 (2008)

Statistic of Natural Images: Scaling in the Woods

D. L. Rudeman, W. Bialek

PRL **73** 814 (1994)