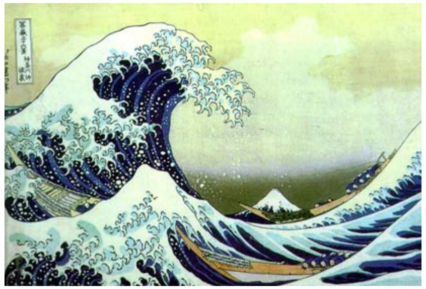


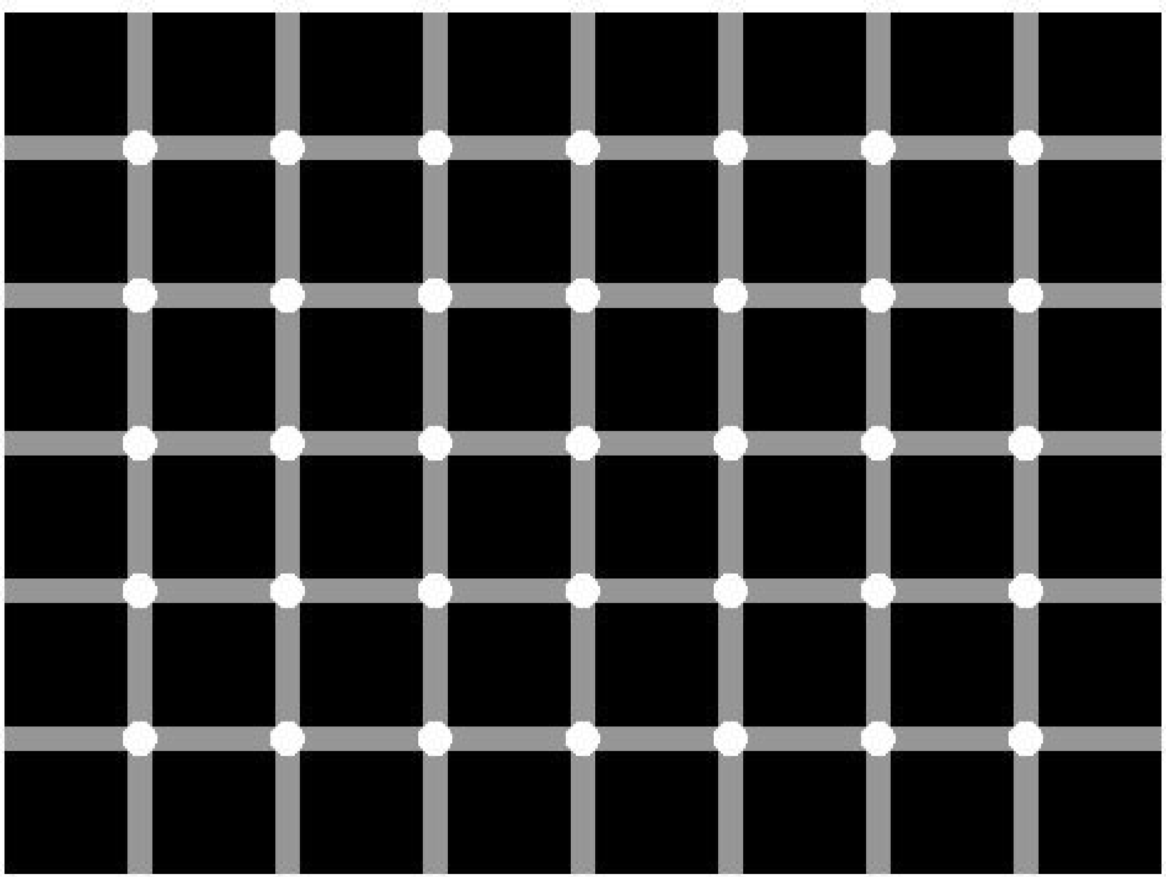
Multimedia communications

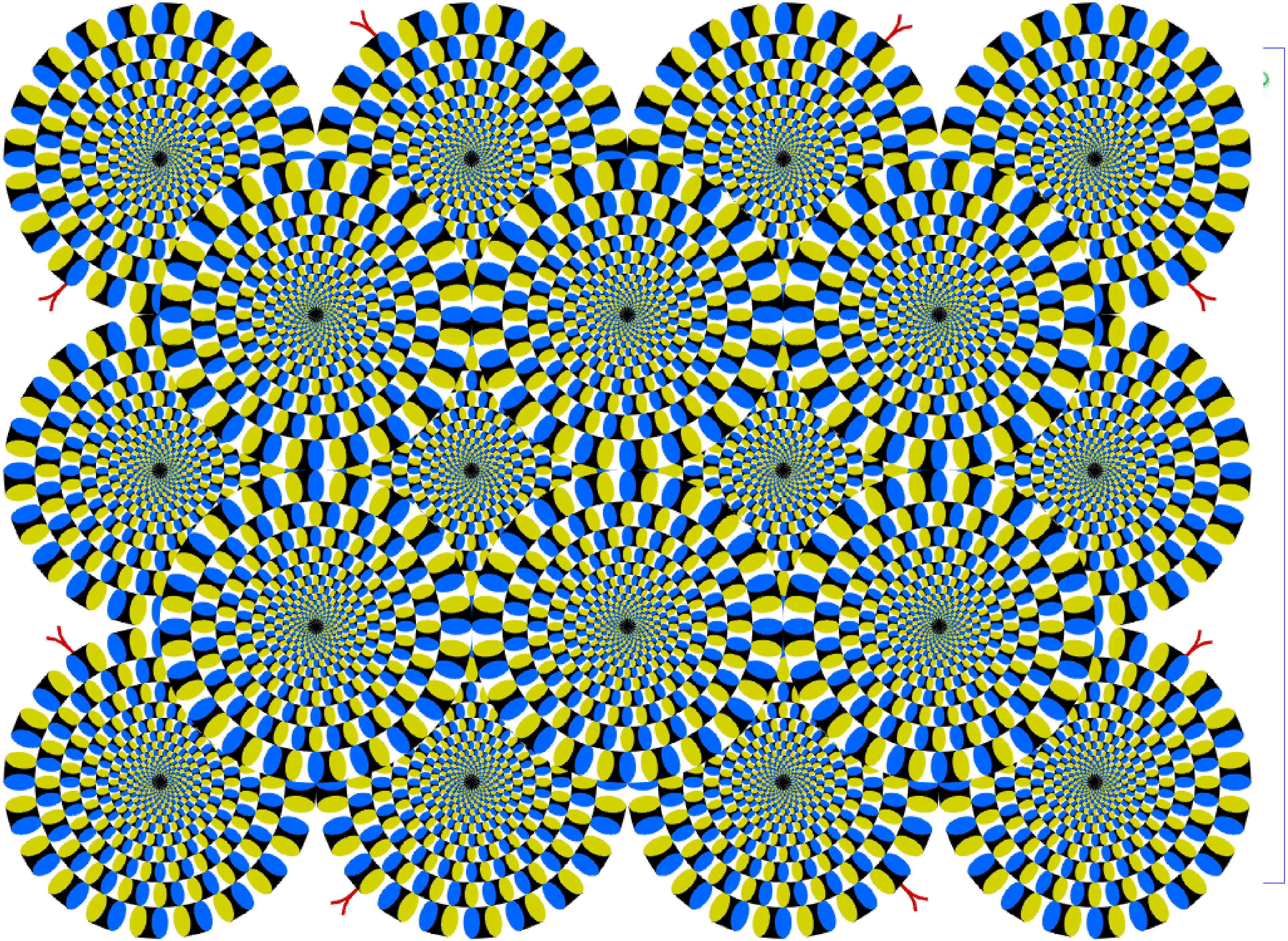
Comunicazione multimediale

G. Menegaz

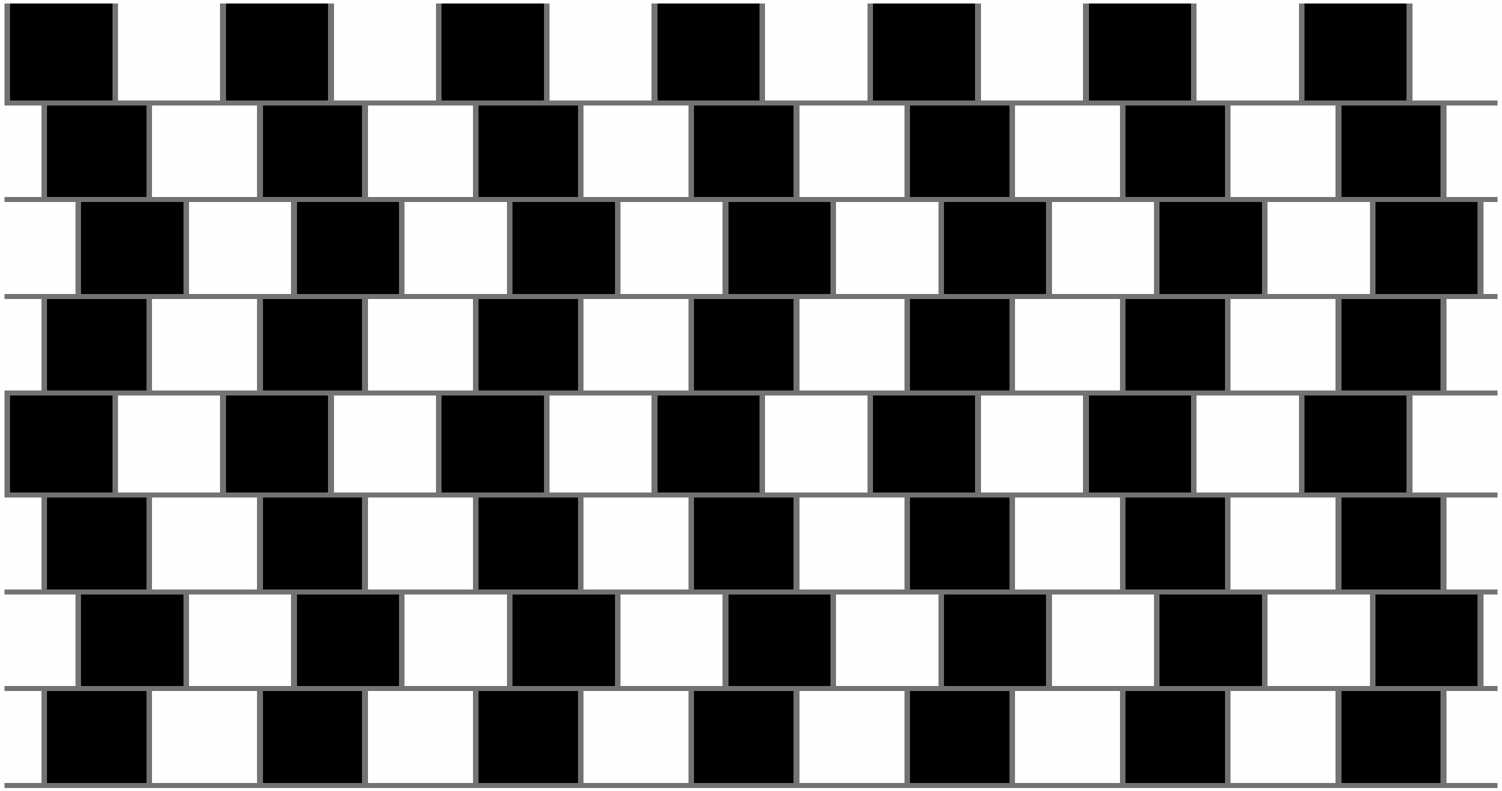
gloria.menegaz@univr.it

Prologue

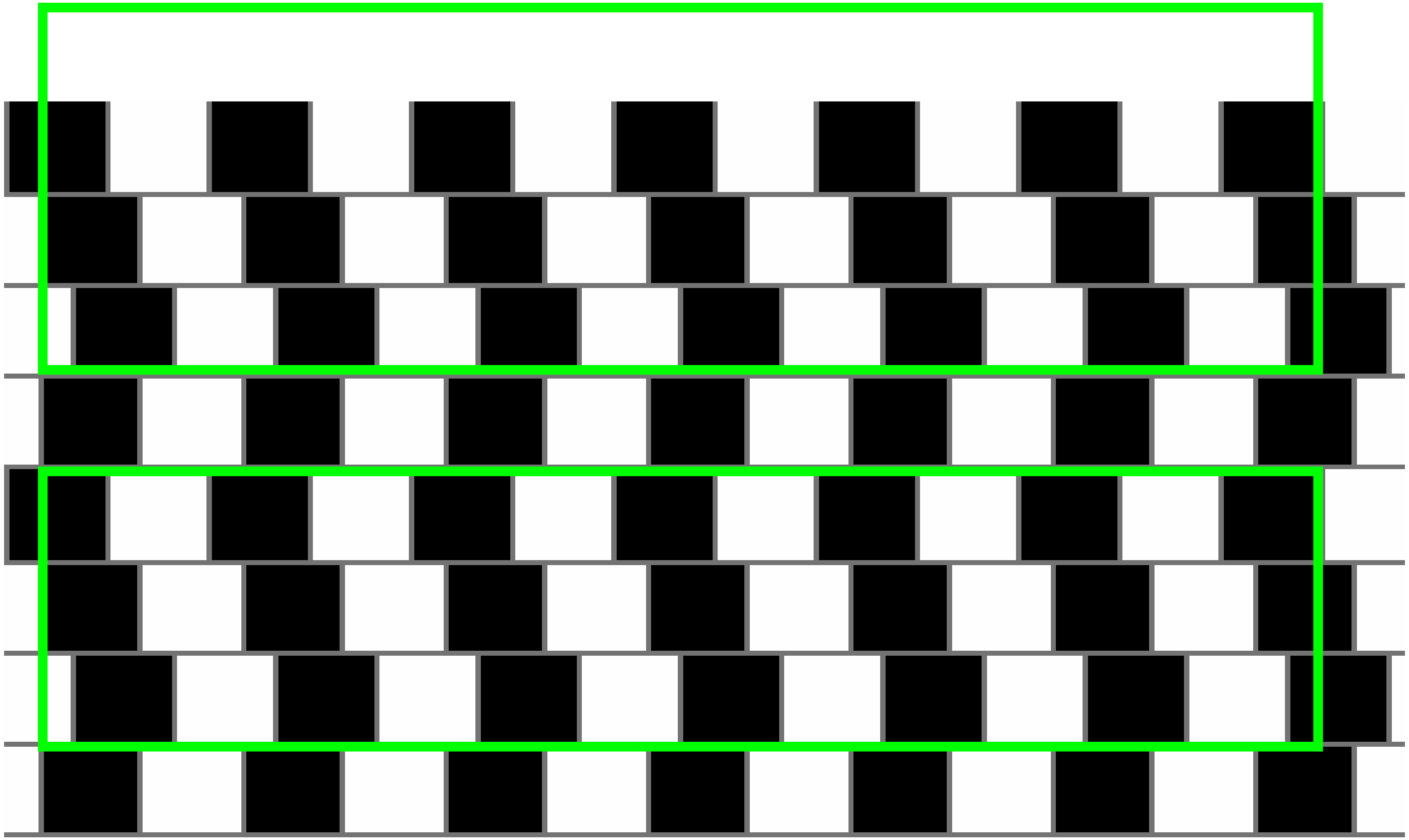




“Context”



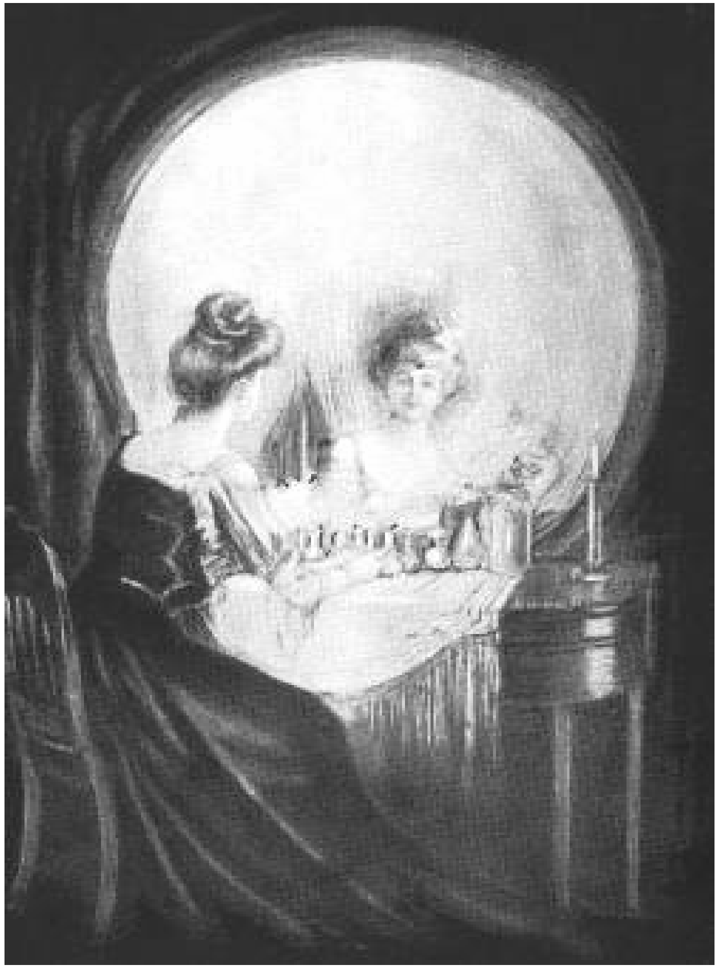
“Context”



“Scale”

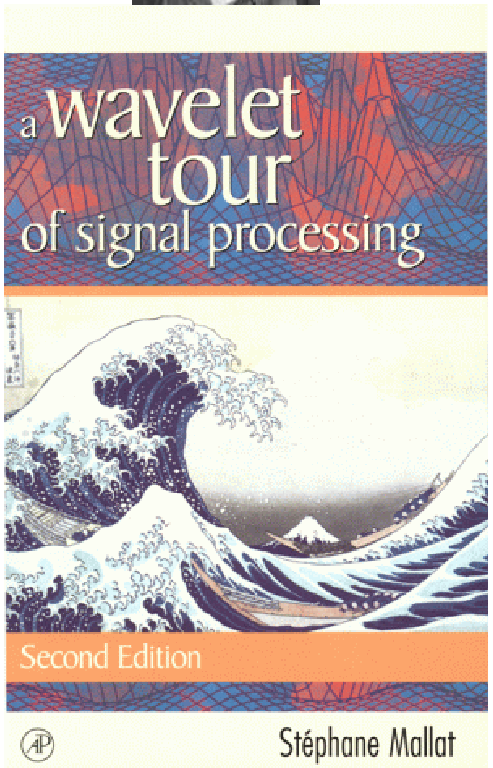
“Scale”

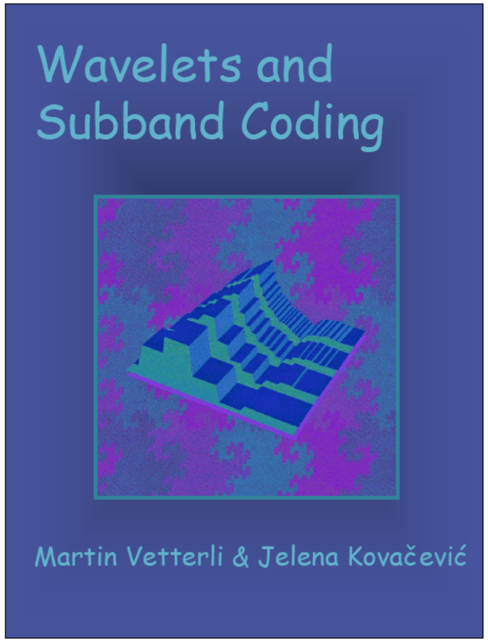
“Scale”



Course overview

- Goal
  - The course is about wavelets and multiresolution
    - Theory: 4 hours per week
      - Tuesday 10.30-13.30 (10.30-12.00 + 12.45-13.30)
      - Wed. 9.30-10.30
    - Laboratory
      - Wed. 14.30-16.30 (Lab. Gamma)
- Contents
  - Review of Fourier theory
  - Wavelets and multiresolution
  - Review of Information theoretic concepts
  - Applications
    - Image coding (JPEG2000)
    - Feature extraction and signal/image analysis

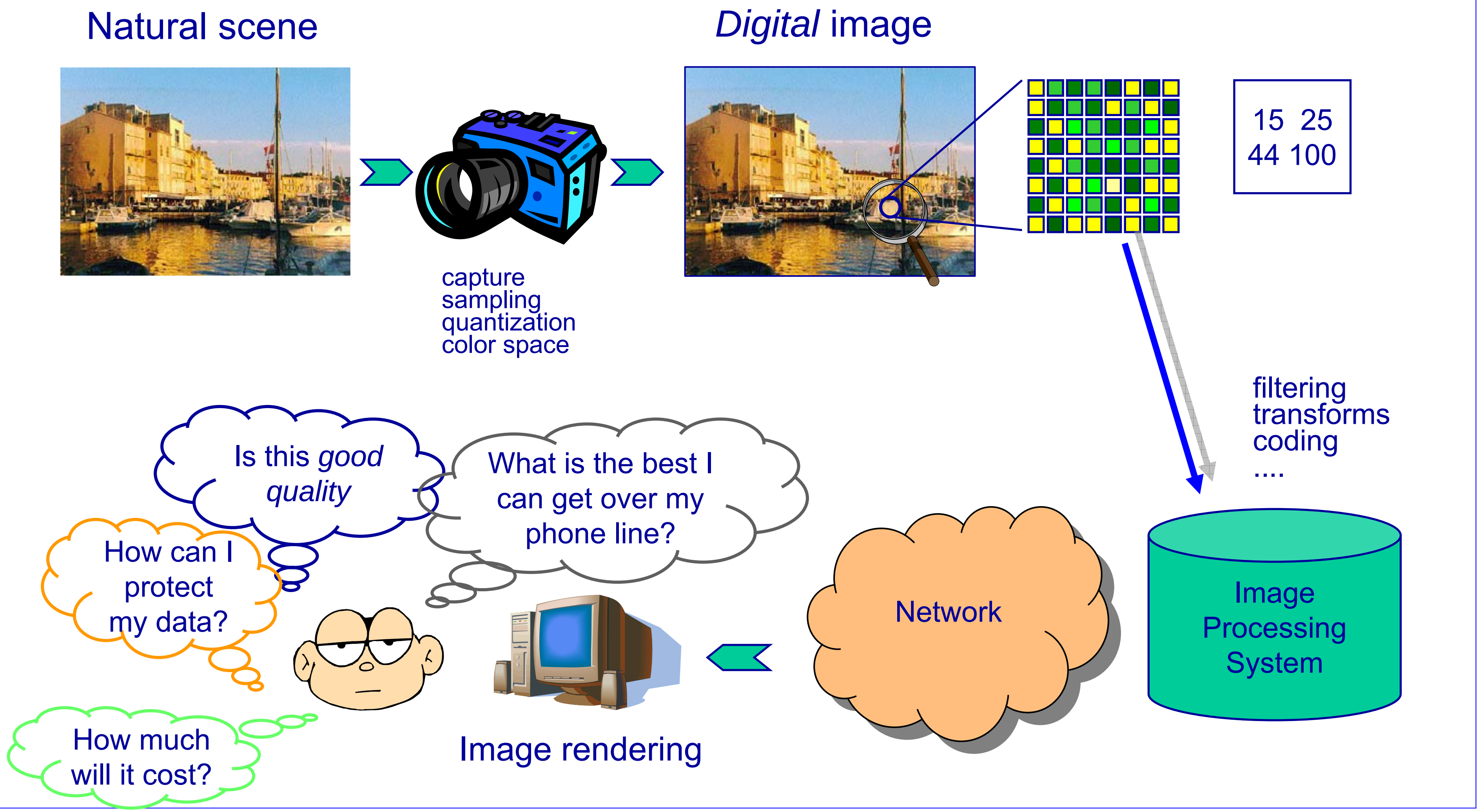




Next 2 weeks scheduling

- Tue. March 9: moved to Thu. March 11 14.30-16.30, Room B
- Wed. March 17: moved from 10.30-11.30 to 16.30-17.30 (or 18.30), Room H

Multimedia framework



Telecommunications for Multimedia

Good news

- It is fun!
- Get in touch with the state-of-the-art technology
- Convince yourself that the time spent on maths&stats was not wasted
- Learn how to map theories into applications
- Acquiring the tools for doing good research!

Bad news

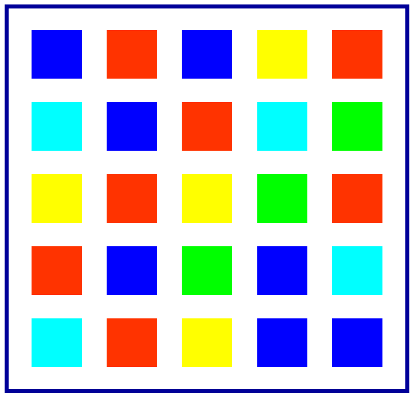
- Some theoretical background is unavoidable
  - Mathematics
    - Fourier transform
    - Linear operators
    - Digital filters
    - Wavelet transform
  - (some) Information theory

Issues in mulitmedia systems

- Broadcasting needs high information carrying capacity
  - Efficient data representation
    - Projection into suitable (perception based?) spaces
    - Color processing
  - Efficient encoding
    - Reduction of redundancy
      - Classical information theoretical principles (entropy based)
      - Novel approaches based on visual perception (perception based)
- Standardization
  - Openness
    - Ability to adapt to new technologies
  - Flexibility
    - Ability to interact with different media
  - JPEG2000, MPEG4, MPEG7

Digital images acquisition

- Analog camera+A/D converter
- Digital cameras
  - CCDs (Charge Coupled Devices)
  - CMOS technology
- In both cases: optics
  - lenses, diaphrams

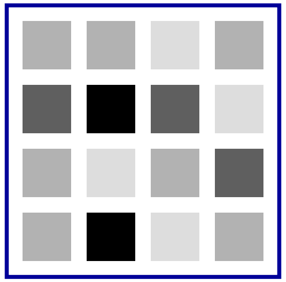


Matrices of photo sensors collecting photons of given wavelength

Features of the capture devices:

- Size and number of photosites
- Noise
- Transfer function of the optical filter

Basics: graylevel images



100	100	200	90
50	0	50	200
100	200	100	50
100	0	200	100

Images : Matrices of numbers
Image processing : Operations among numbers
bit depth : number of bits/pixel
N bit/pixel : 2^{N-1} shades of gray (typically N=8)

Sampling

- Sampling in p-dimensions

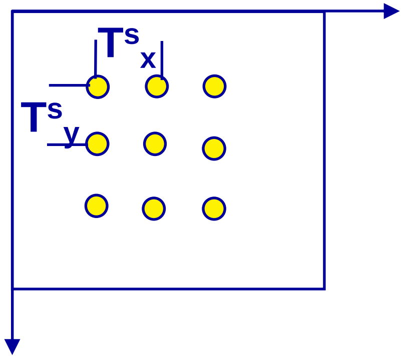
$$s_T(\vec{x}) = \sum_{k \in Z^p} \delta(\vec{x} - kT)$$

$$f_T(\vec{x}) = f(\vec{x}) s_T(\vec{x})$$

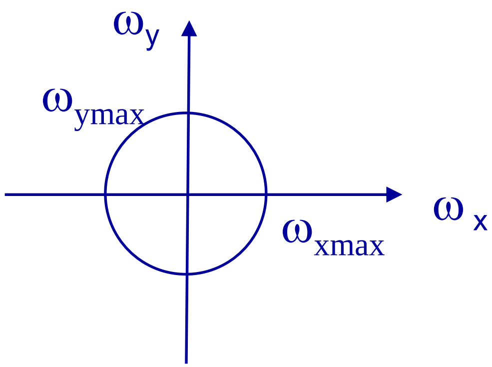
- Nyquist theorem

$$\begin{cases} \omega_x^s \geq 2\omega_{x\max} \\ \omega_y^s \geq 2\omega_{y\max} \end{cases} \Rightarrow \begin{cases} T_x^s \leq 2\pi \dfrac{1}{2\omega_{x\max}} \\ T_y^s \leq 2\pi \dfrac{1}{2\omega_{y\max}} \end{cases}$$

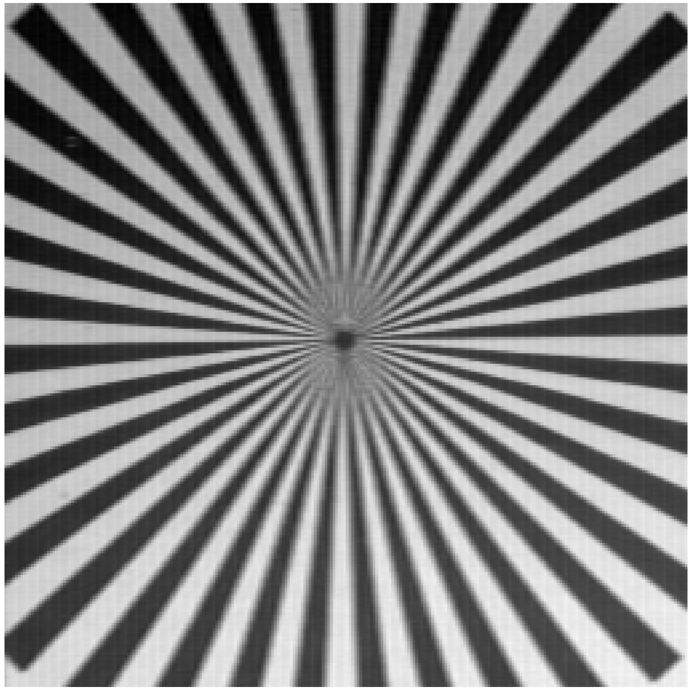
2D spatial domain

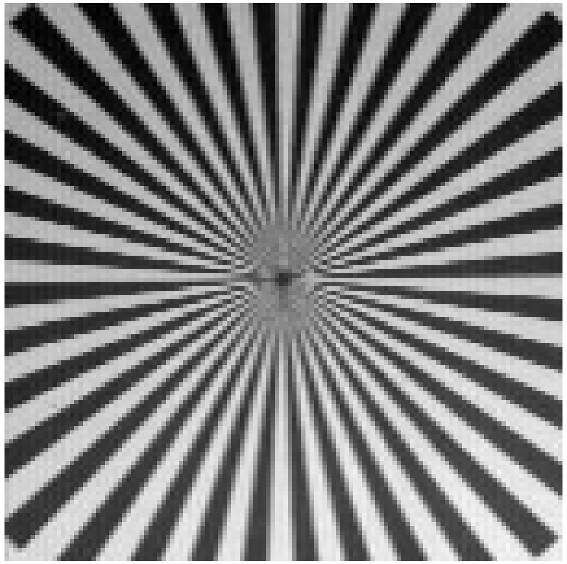


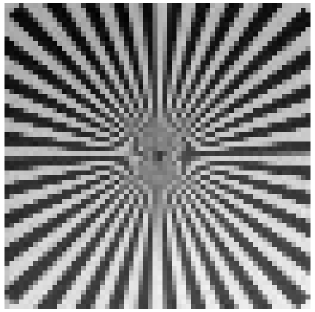
2D Fourier domain



Spatial aliasing

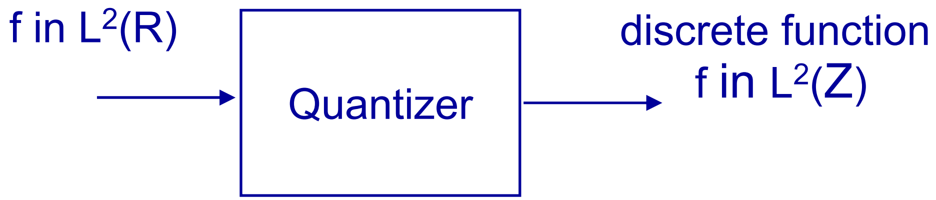


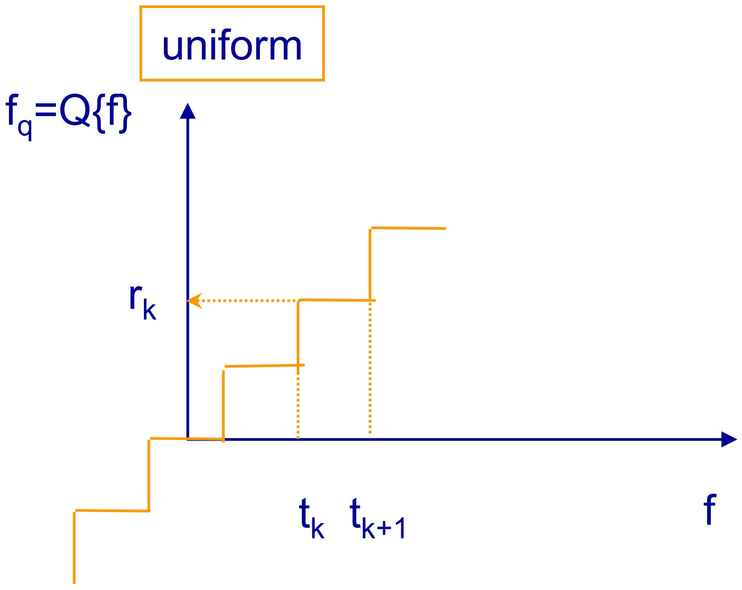


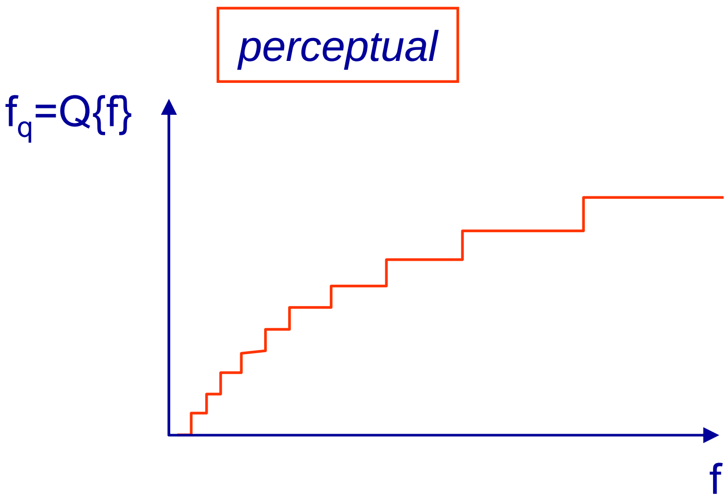


Quantization

- A/D conversion ⇨ quantization



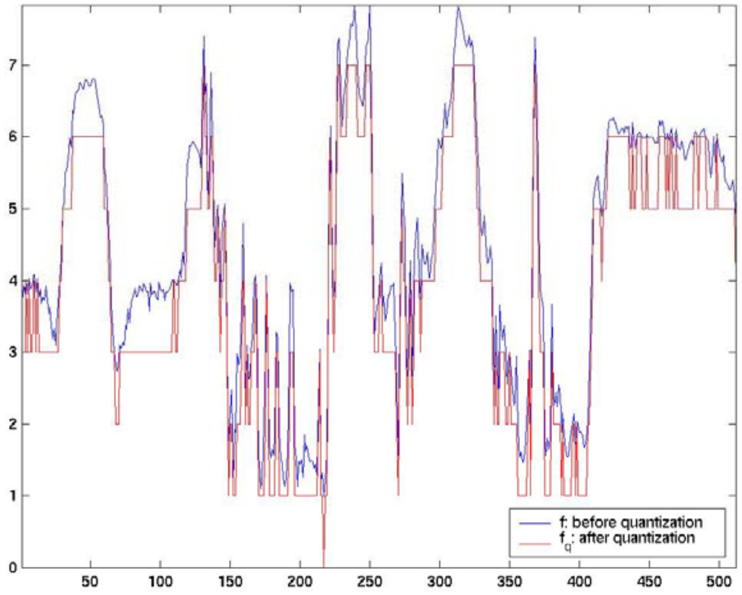




The sensitivity of the eye decreases increasing the background intensity (Weber law)

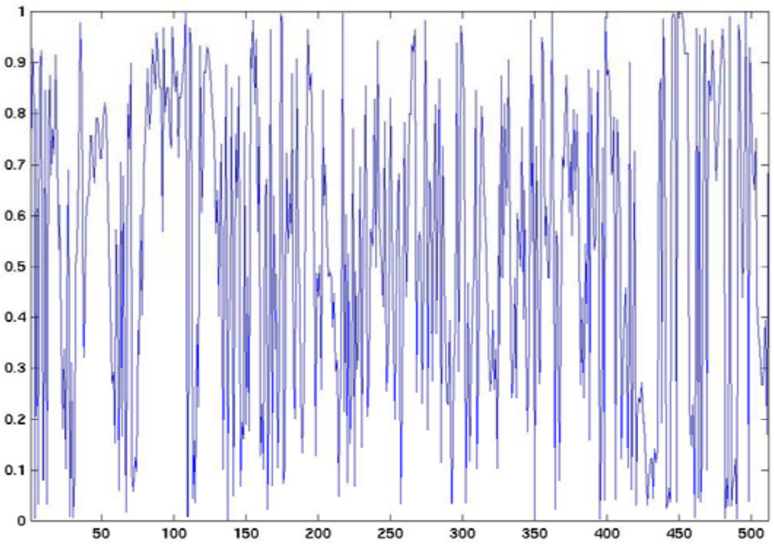
Quantization

Signal before (blue) and after quantization (red) Q



Equivalent noise: $n=f_q - f$

additive noise model: $f_q=f+n$



Distortion measure

- Distortion measure

$$D = \mathrm{E}\left[\left(f_Q - f\right)^2\right] = \sum_{k=0}^{K} \int_{t_k}^{t_{k+1}} \left(f_Q - f\right)^2 p(f)\,df$$

  - The distortion is measured as the expectation of the mean square error (MSE) difference between the original and quantized signals.

- Lack of correlation with perceived image quality
  - Even though this is a very natural way for the quantification of the quantization artifacts, it is not representative of the *visual annoyance* due to the majority of common artifacts.

Quantization

original

5 levels

10 levels

50 levels

Mathematical tools

Introduction

- Sparse representations: few coefficients reveal the information we are looking for.
  - Such representations can be constructed by decomposing signals over elementary waveforms chosen in a family called a *dictionary*.
  - An orthogonal basis is a dictionary of minimum size that can yield a sparse representation if designed to concentrate the signal energy over a set of few vectors. This set gives a *geometric* signal description.
    - Signal compression and noise reduction
  - Dictionaries of vectors that are larger than bases are needed to build sparse representations of complex signals. But choosing is difficult and requires more complex algorithms.
    - Sparse representations in redundant dictionaries can improve pattern recognition,compression,and noise reduction
- Basic ingredients: Fourier and Wavelet transforms
  - They decompose signals over oscillatory waveforms that reveal many signal properties and provide a path to sparse representations

Signals as functions

- CT analogue signals (real valued functions of continuous independent variables)
  - 1D: $f=f(t)$
  - 2D: $f=f(x,y)$ x,y
  - Real world signals (audio, ECG, pictures taken with an analog camera)
- DT analogue signals (real valued functions of discrete variables)
  - 1D: $f=f[k]$
  - 2D: $f=f[i,j]$
  - *Sampled* signals
- Digital signals (discrete valued functions of DT variables)
  - 1D: $y=y[k]$
  - 2D: $y=y[i,j]$
  - *Sampled and discretized* signals

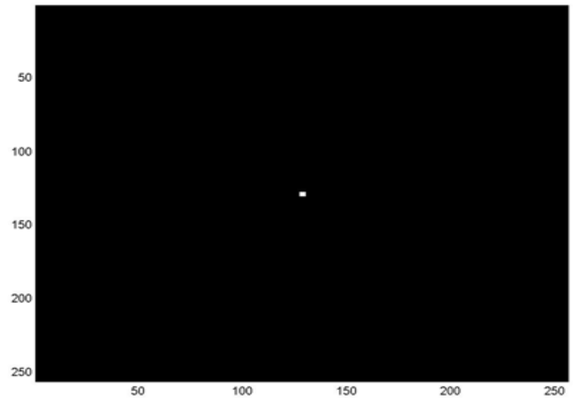
Images as functions

- Gray scale images: 2D functions
  - Domain of the functions: set of *(x,y)* values for which *f(x,y)* is defined : 2D lattice *[i,j]* defining the pixel locations
  - Set of values taken by the function : gray levels

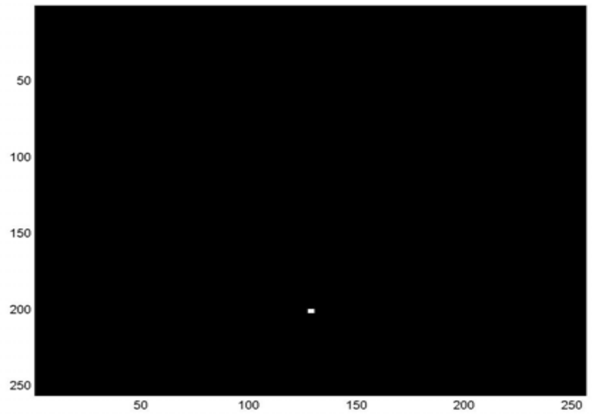
- Digital images can be seen as functions defined over a discrete domain $\{i,j: 0<i<I, 0<j<J\}$
  - *I,J:* number of rows (columns) of the matrix corresponding to the image
  - *f=f[i,j]:* gray level in position *[i,j]*

Example 1: δ function

$$\delta[i,j] = \begin{cases} 1 & i = j = 0 \\ 0 & i,j \neq 0; i \neq j \end{cases}$$



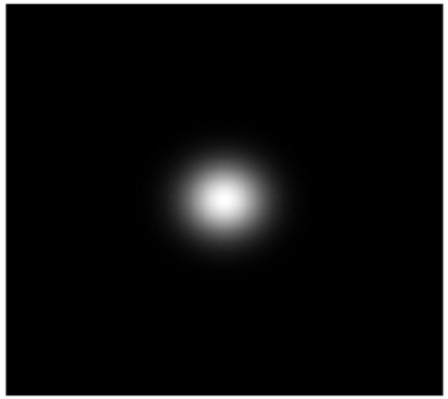
$$\delta[i, j - J] = \begin{cases} 1 & i = 0; j = J \\ 0 & otherwise \end{cases}$$



Example 2: Gaussian

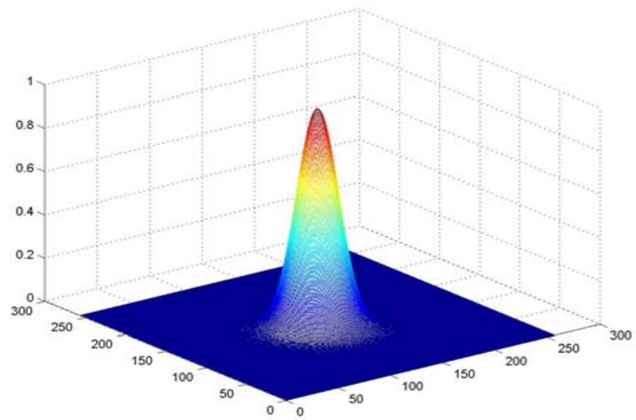
Continuous function

$$f(x,y)=\frac{1}{\sigma\sqrt{2\pi}}e^{\frac{x^2+y^2}{2\sigma^2}}$$

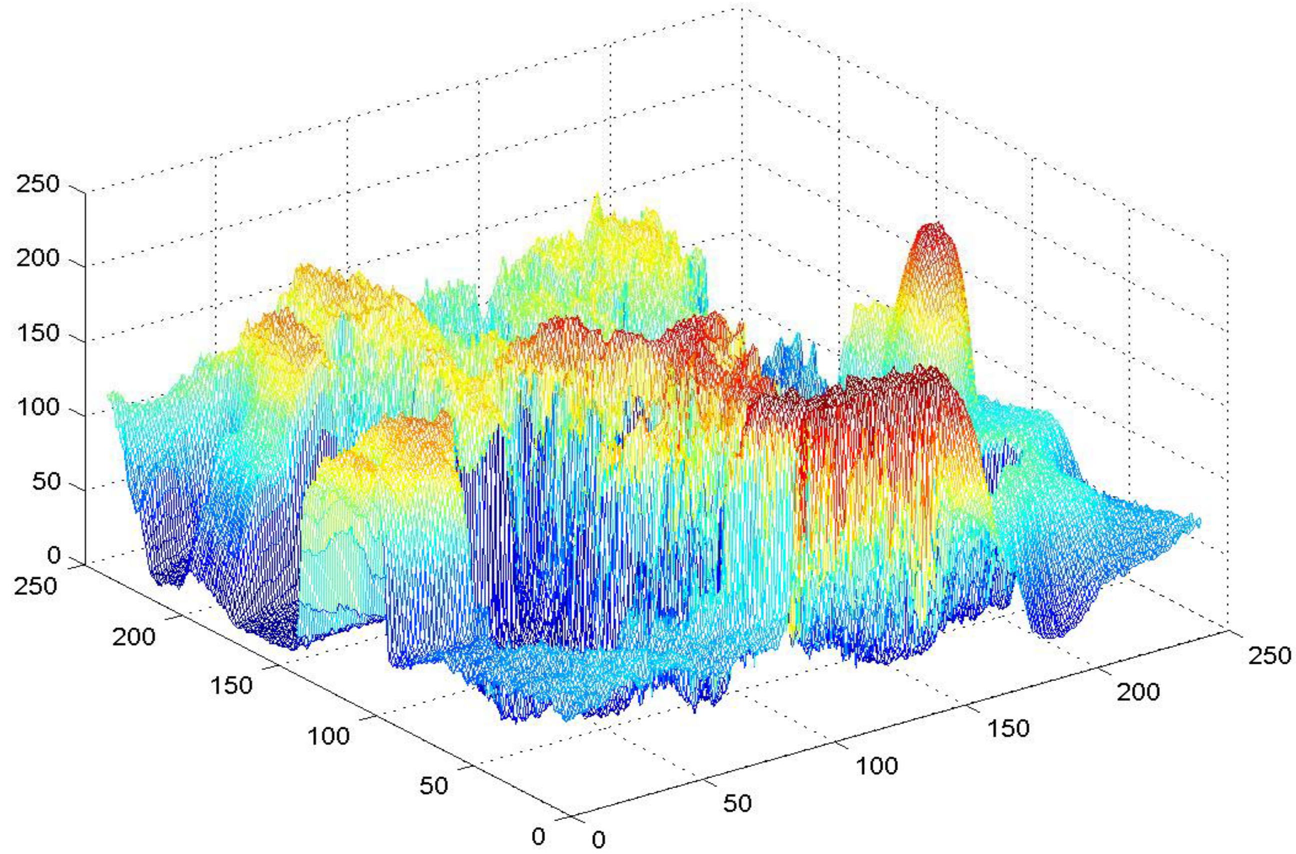


Discrete version

$$f[i,j]=\frac{1}{\sigma\sqrt{2\pi}}e^{\frac{i^2+j^2}{2\sigma^2}}$$



Example 3: Natural image



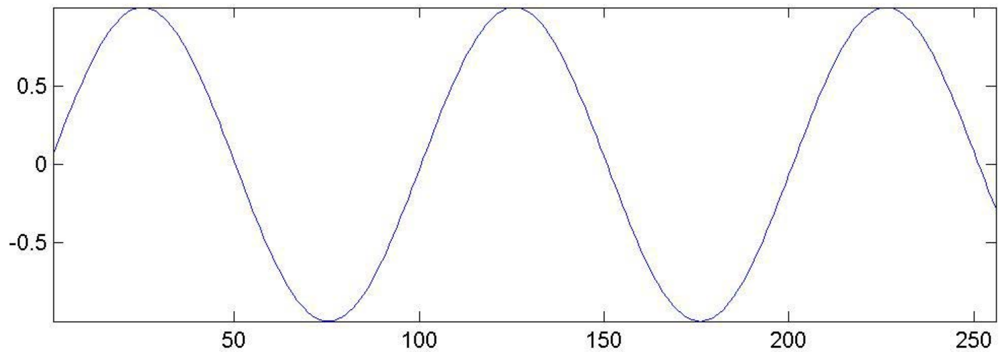
Example 3: Natural image

The Fourier kingdom

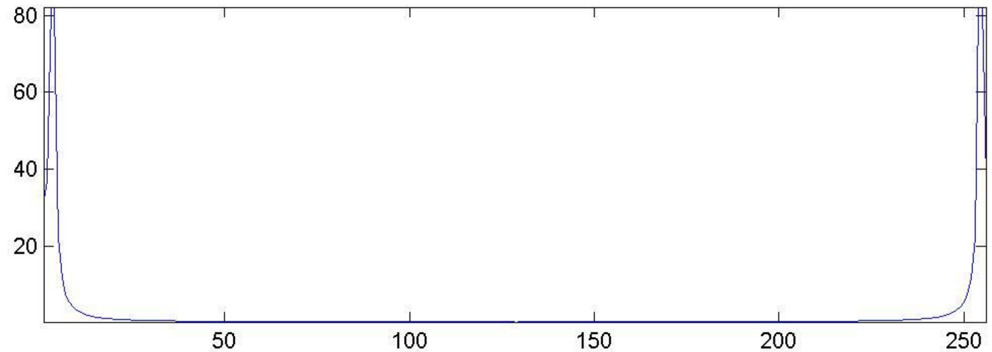
- Frequency domain characterization of signals

$$F(\omega)=\int_{-\infty}^{+\infty} f(t)e^{-j\omega t}dt$$

$$f(t)=\int_{-\infty}^{+\infty} F(\omega)e^{j\omega t}dt$$

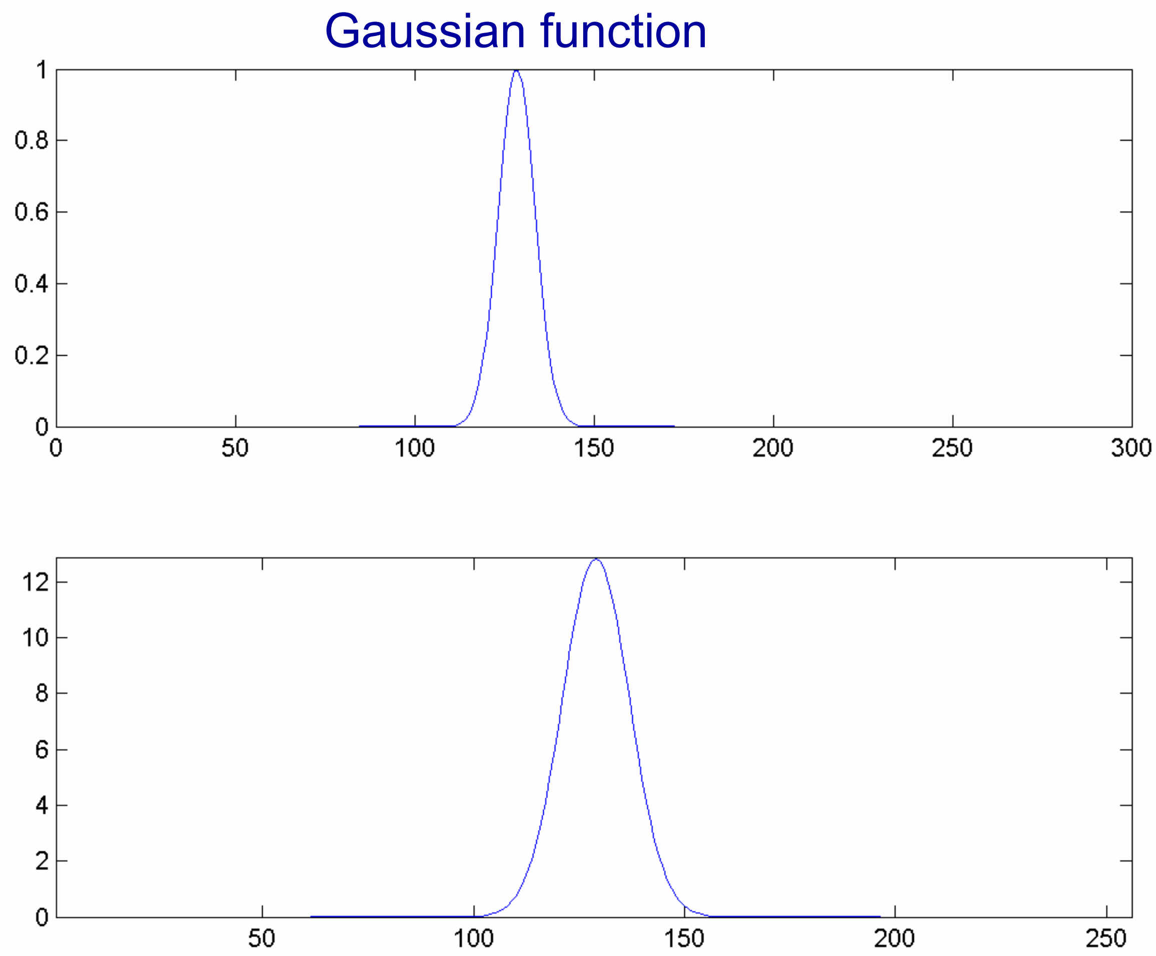


Signal domain



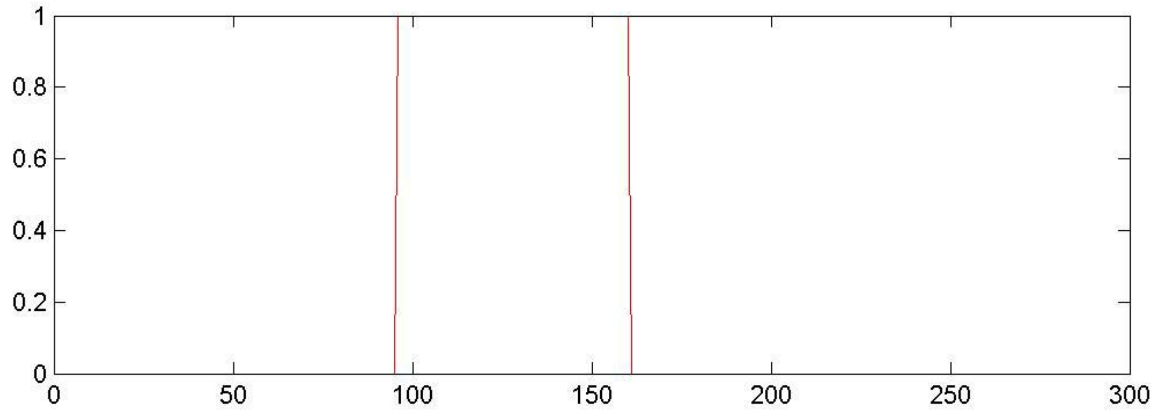
Frequency domain

The Fourier kingdom

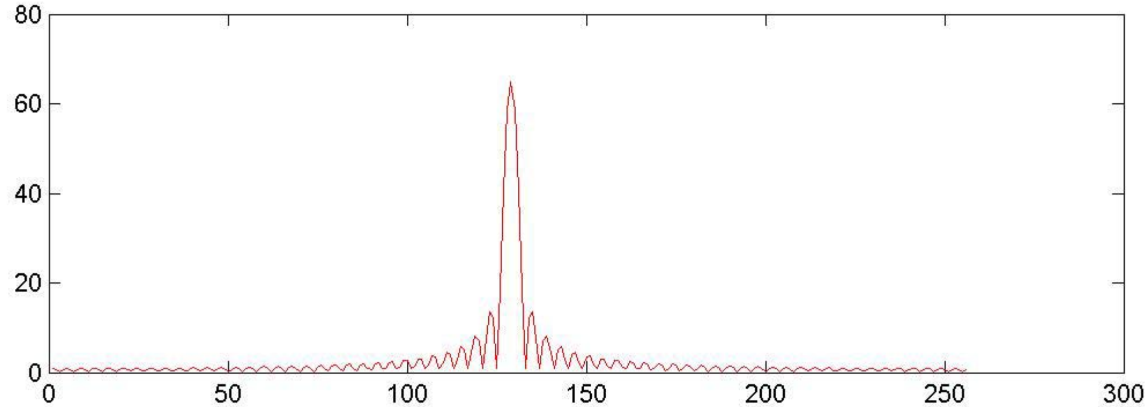


The Fourier kingdom

rect function



sinc function



2D Fourier transform

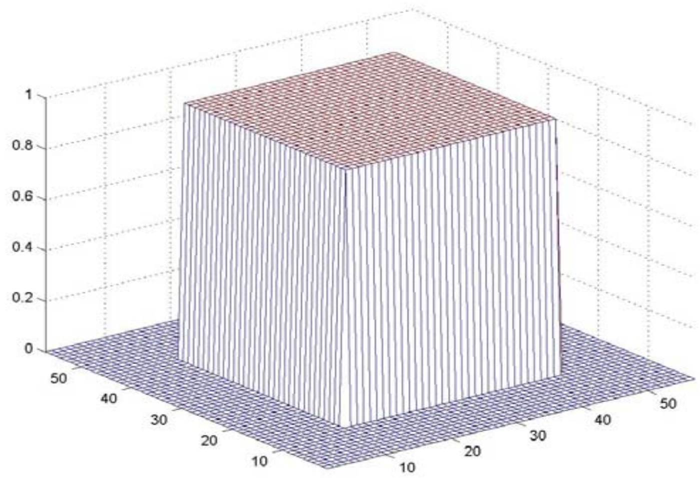
$$\hat{f}\left(\omega_x,\omega_y\right)=\int_{-\infty}^{+\infty} f\left(x,y\right)e^{-j\left(\omega_x x+\omega_y y\right)}dxdy$$

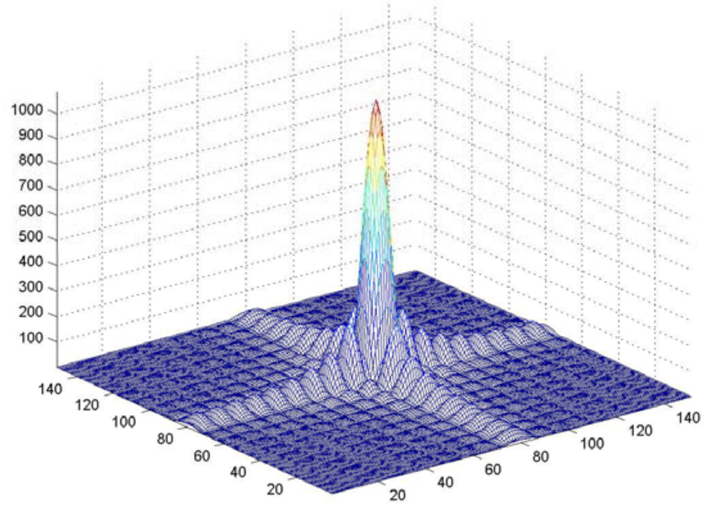
$$f\left(x,y\right)=\frac{1}{4\pi^2}\int_{-\infty}^{+\infty}\hat{f}\left(\omega_x,\omega_y\right)e^{j\left(\omega_x x+\omega_y y\right)}d\omega_x d\omega_y$$

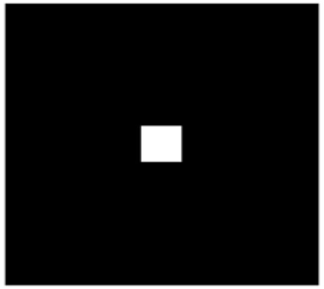
$$\iint f\left(x,y\right)g^*\left(x,y\right)dxdy=\frac{1}{4\pi^2}\iint \hat{f}\left(\omega_x,\omega_y\right)\hat{g}^*\left(\omega_x,\omega_y\right)d\omega_x d\omega_y$$ Parseval formula

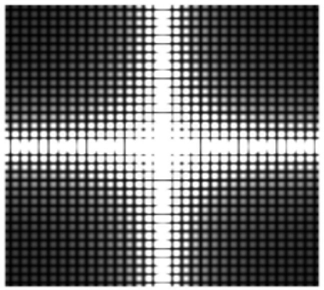
$$f=g\rightarrow\iint\left|f\left(x,y\right)\right|^2 dxdy=\frac{1}{4\pi^2}\iint\left|\hat{f}\left(\omega_x,\omega_y\right)\right|^2 d\omega_x d\omega_y$$ Plancherel equality

The Fourier kingdom

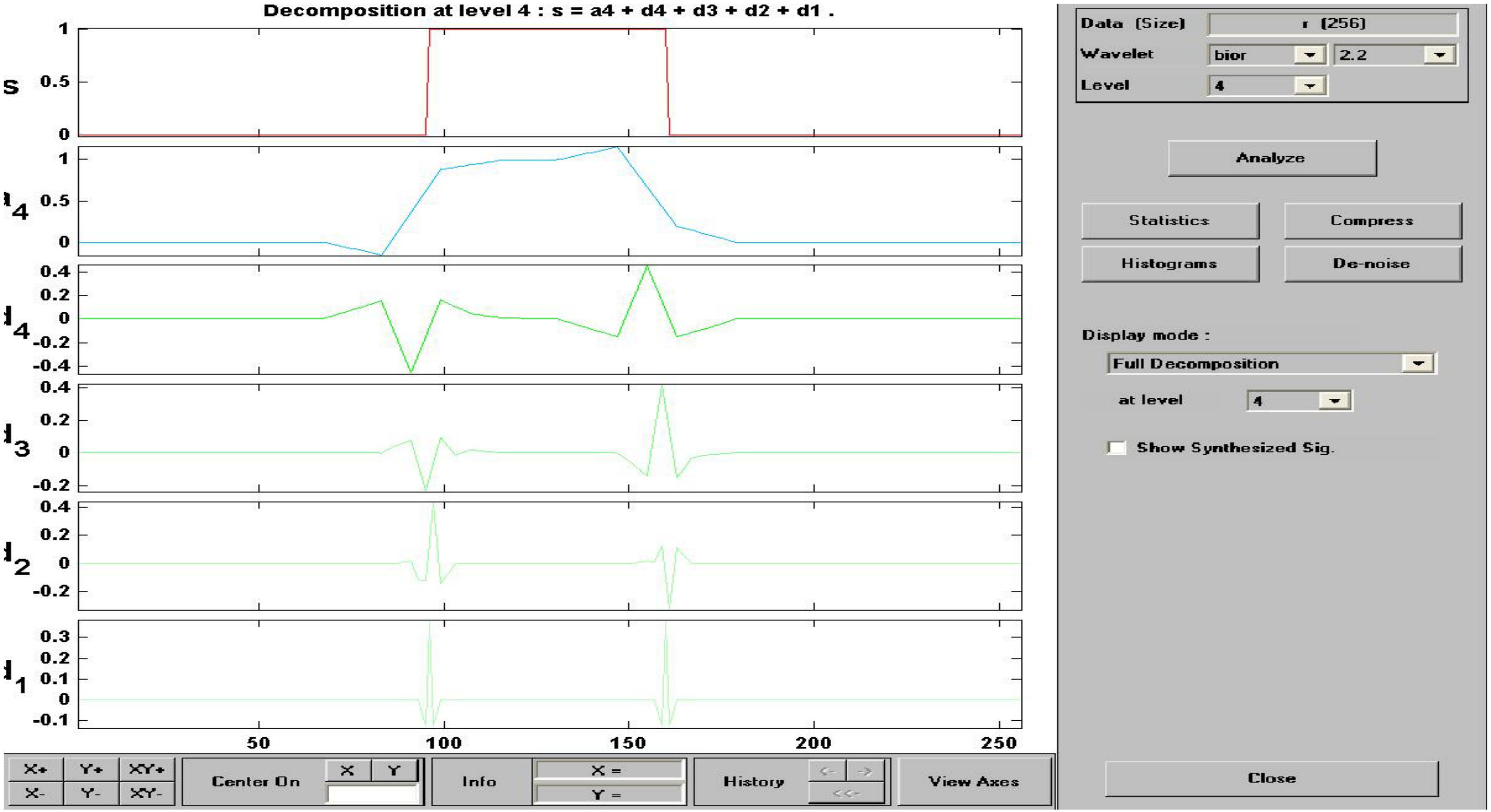




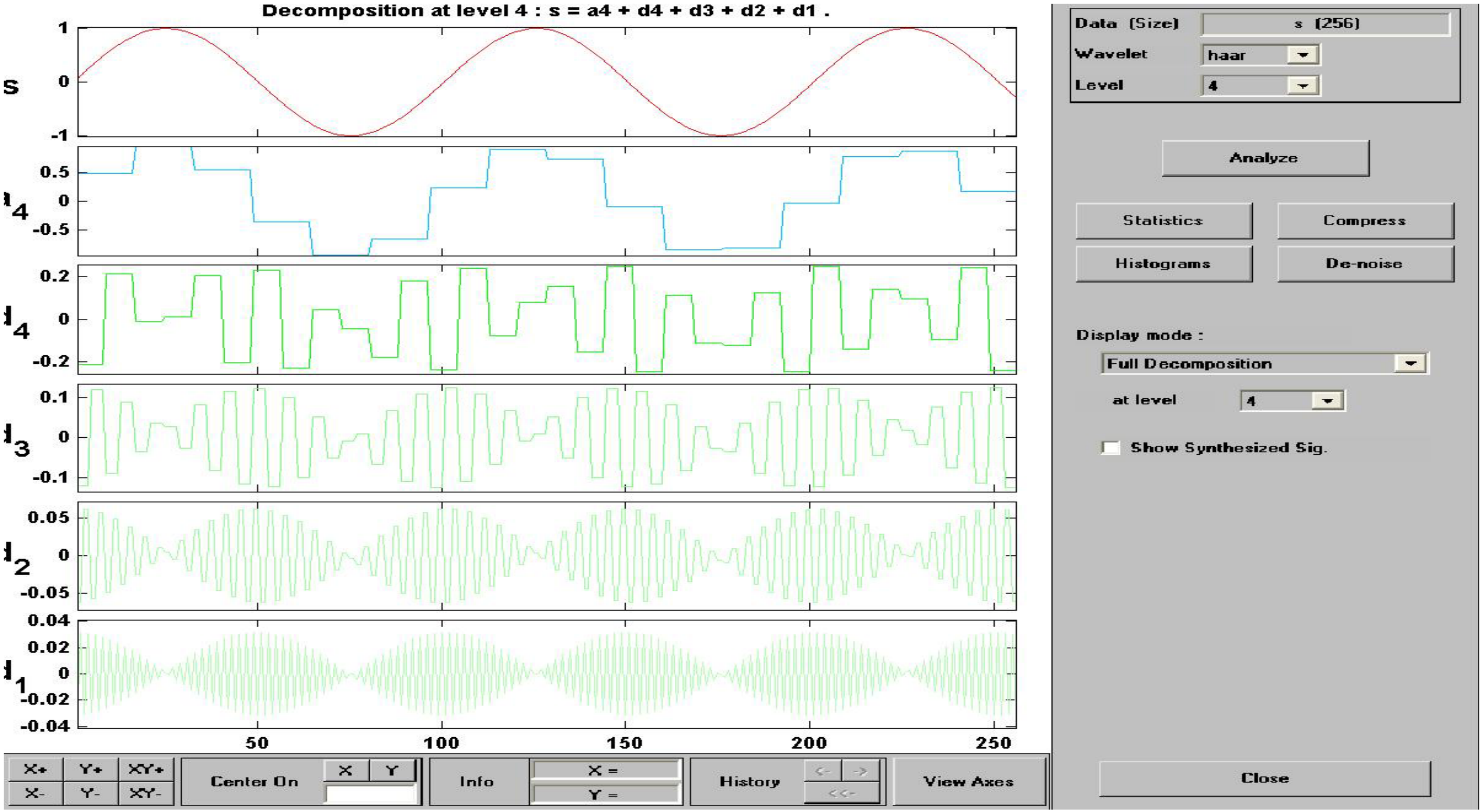




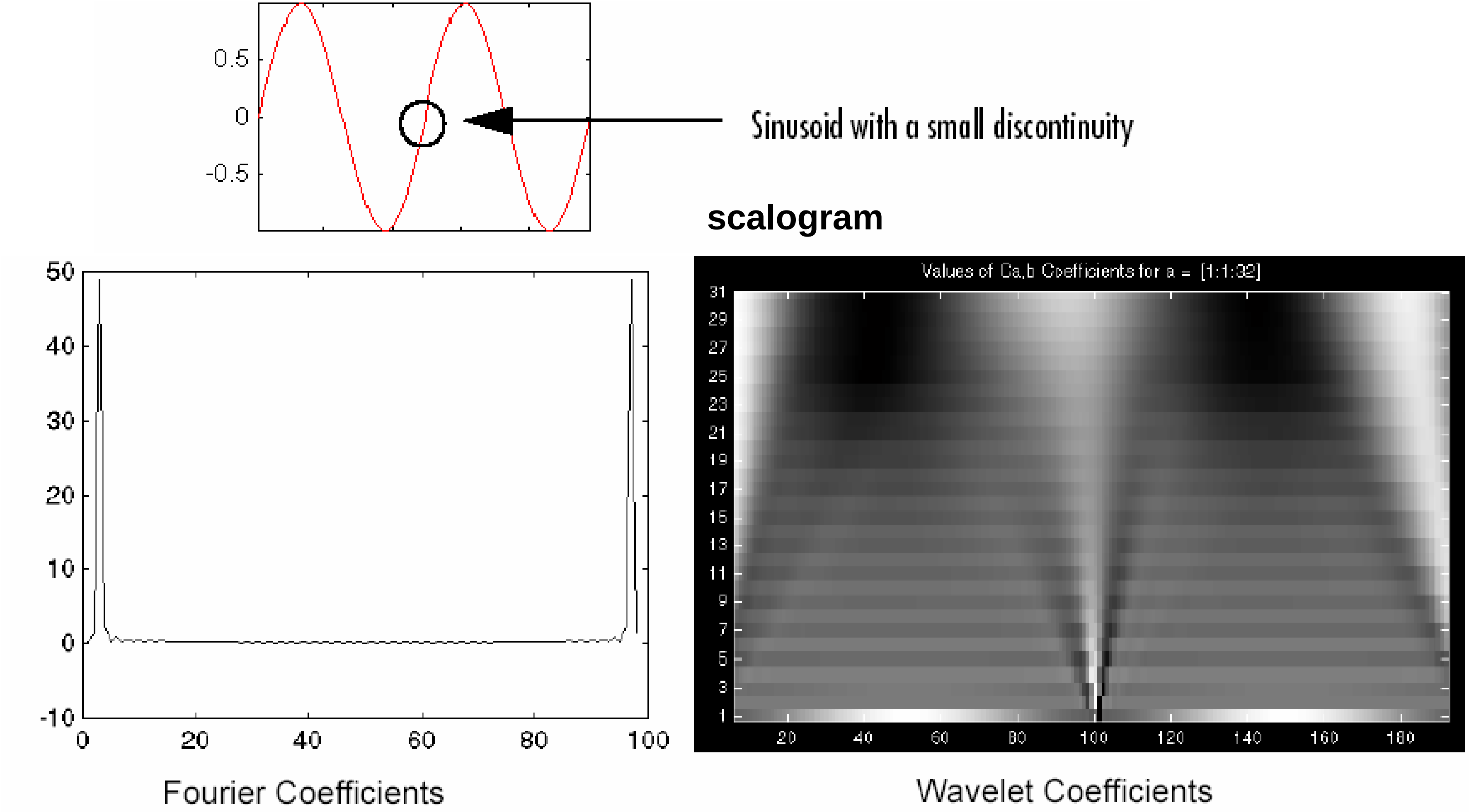
Wavelet representation



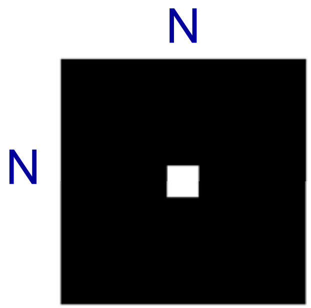
Wavelet representation

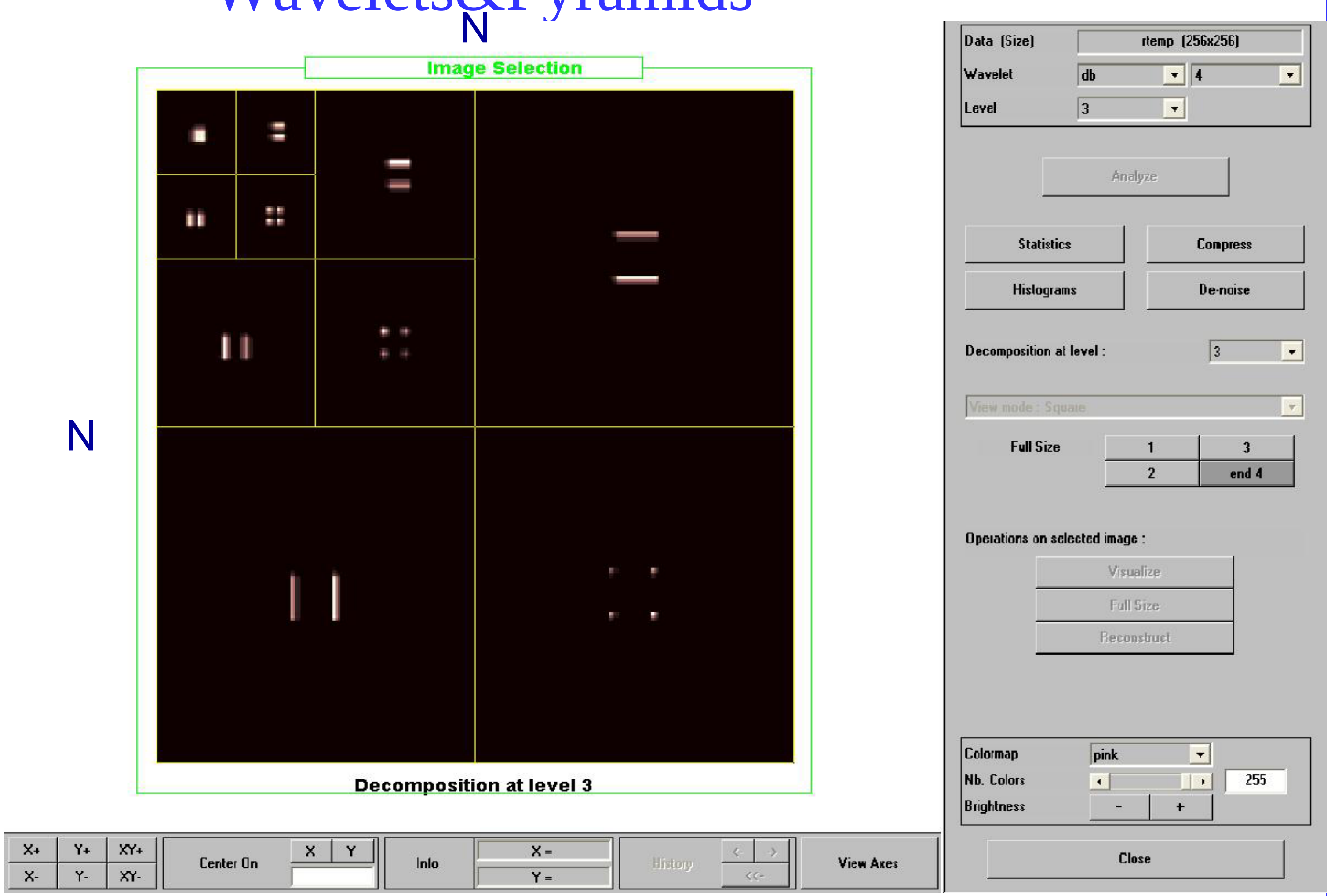


Wavelets are good for transients



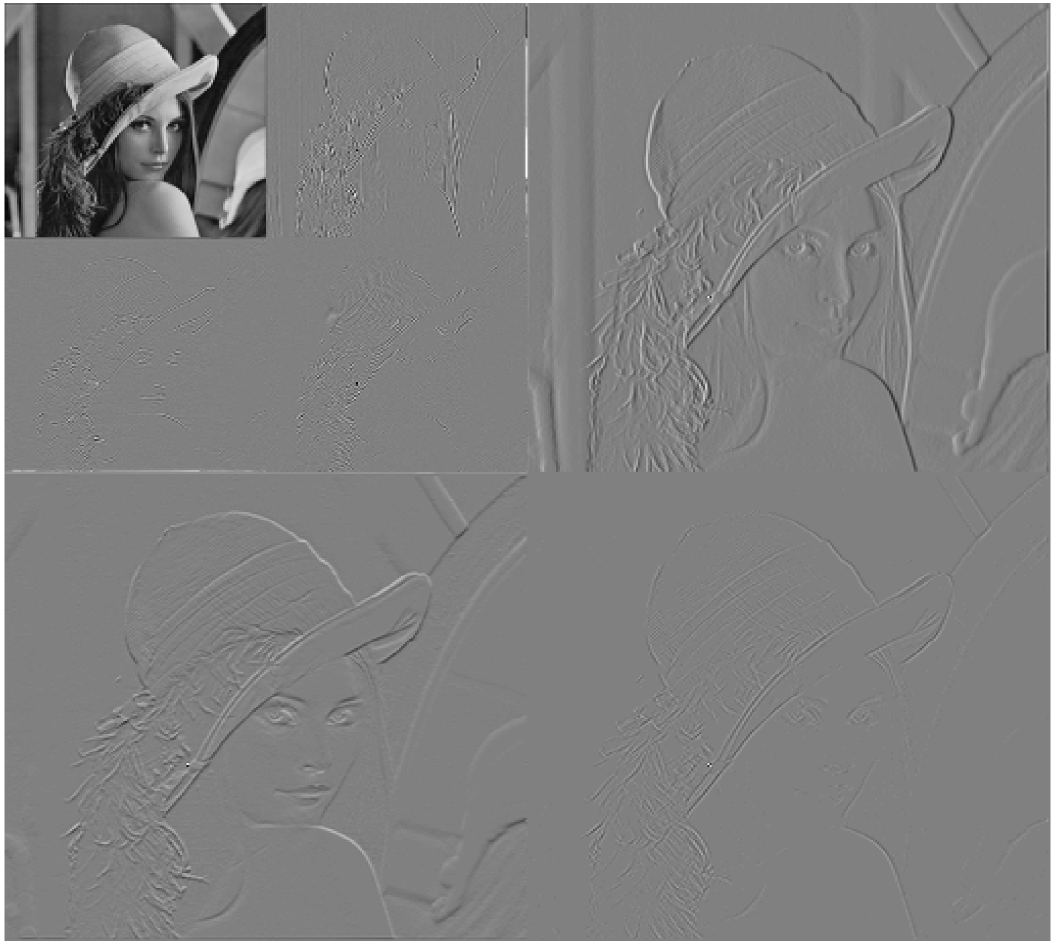
Wavelets&Pyramids



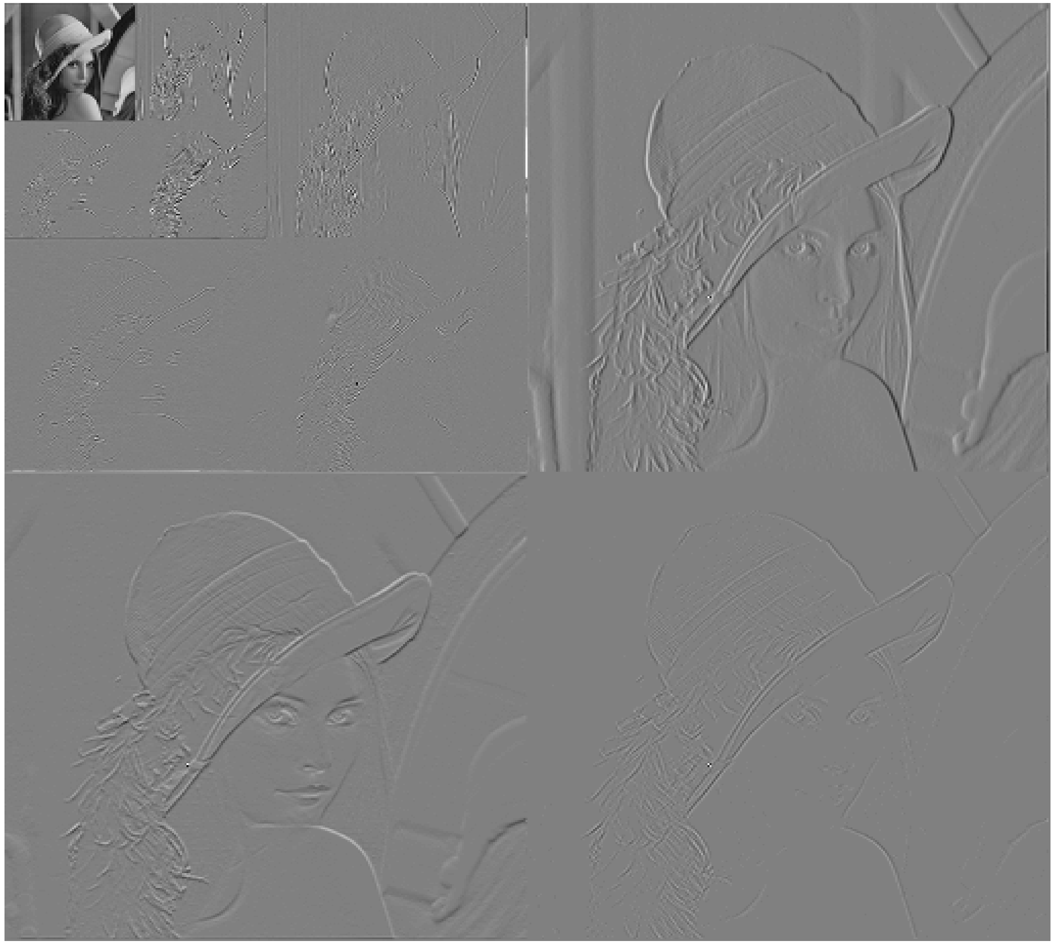


Wavelets&Pyramids

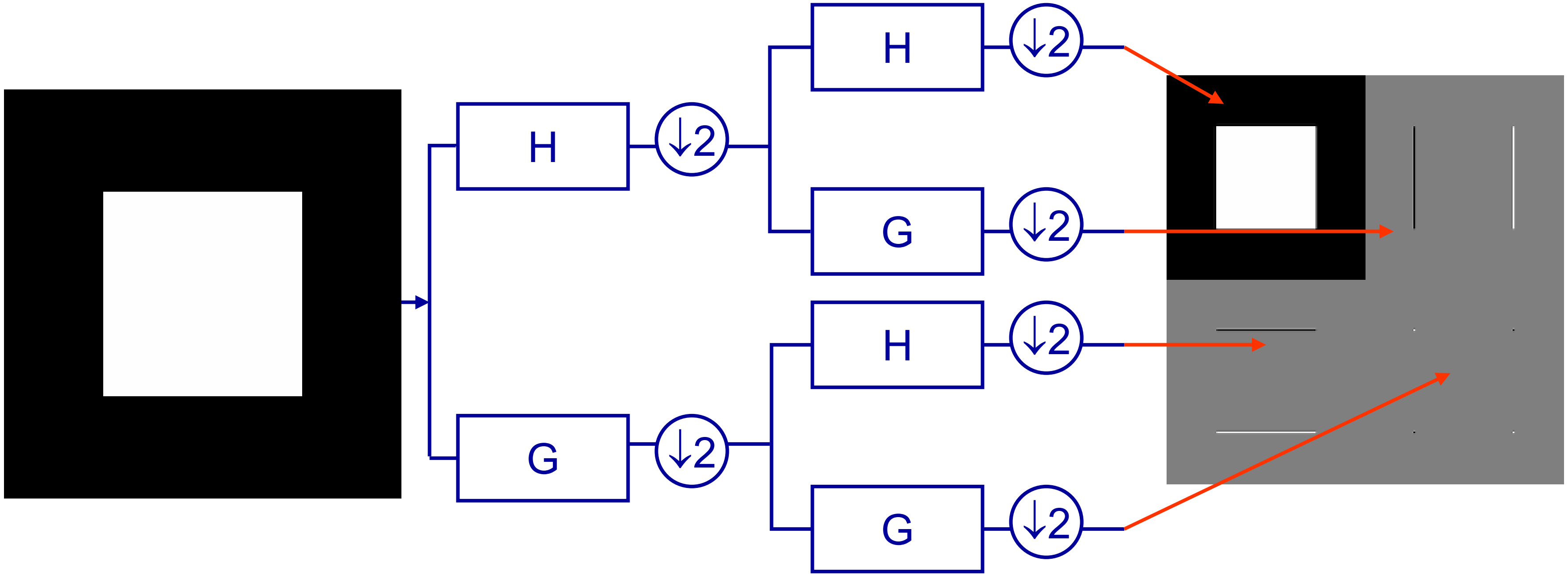
Wavelets&Pyramids



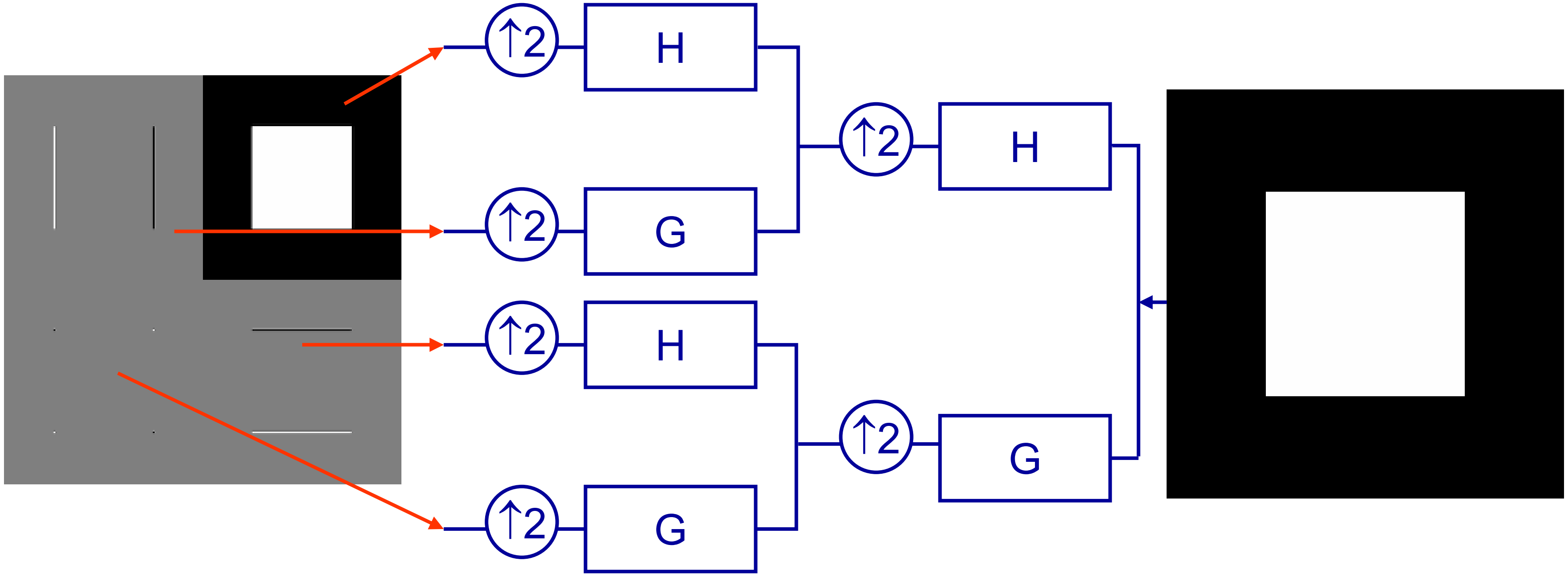
Waveletts&Pyramids



Waveletts&Filterbanks



Waveletts&Filterbanks



Very efficient implementation by recursive filtering

Fourier versus Wavelets

Fourier

- Basis functions are sinusoids
  - More in general, complex exponentials
- Switching from signal domain t to frequency domain f
  - Either spatial or temporal
- Good localization either in time or in frequency
  - Transformed domain: Information on the sharpness of the transient but not on its position
- Good for stationary signals but unsuitable for transient phenomena

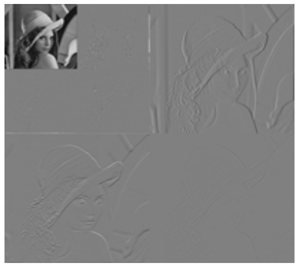
Wavelets

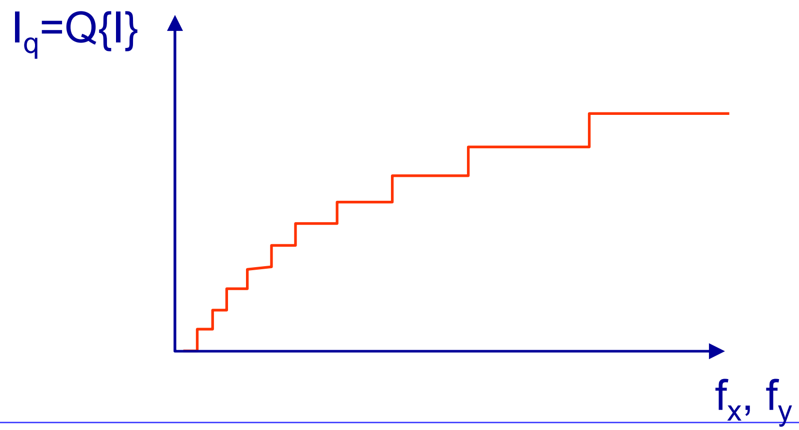
- Different families of basis functions are possible
  - Haar, Daubechies', biorthogonal
- Switching from the signal domain to a *multiresolution* representation
- *Good localization in time and frequency*
  - Information on *both* the *sharpness* of the transient and the *point* where it happens
- Good for any type of signal

Wavelet-based coding

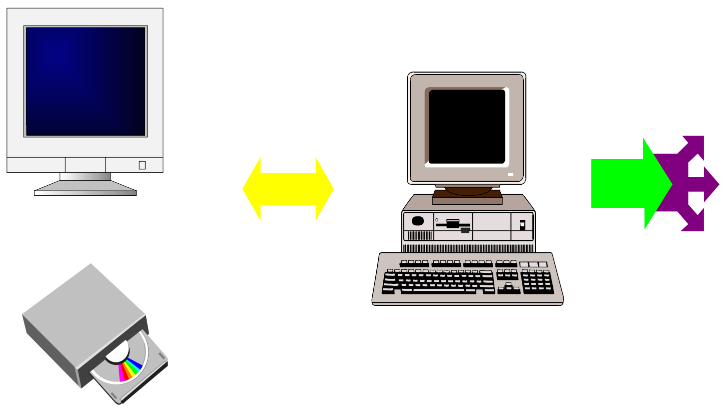
Original image

Wavelet transform (DWT) → (Perceptual) quantization → Entropy coding → Encoded image





Coding standards



Desirable features:

- Flexibility
- User-data interactivity
- Openness
- Easy to use
- User interactivity
- Security

JPEG2000