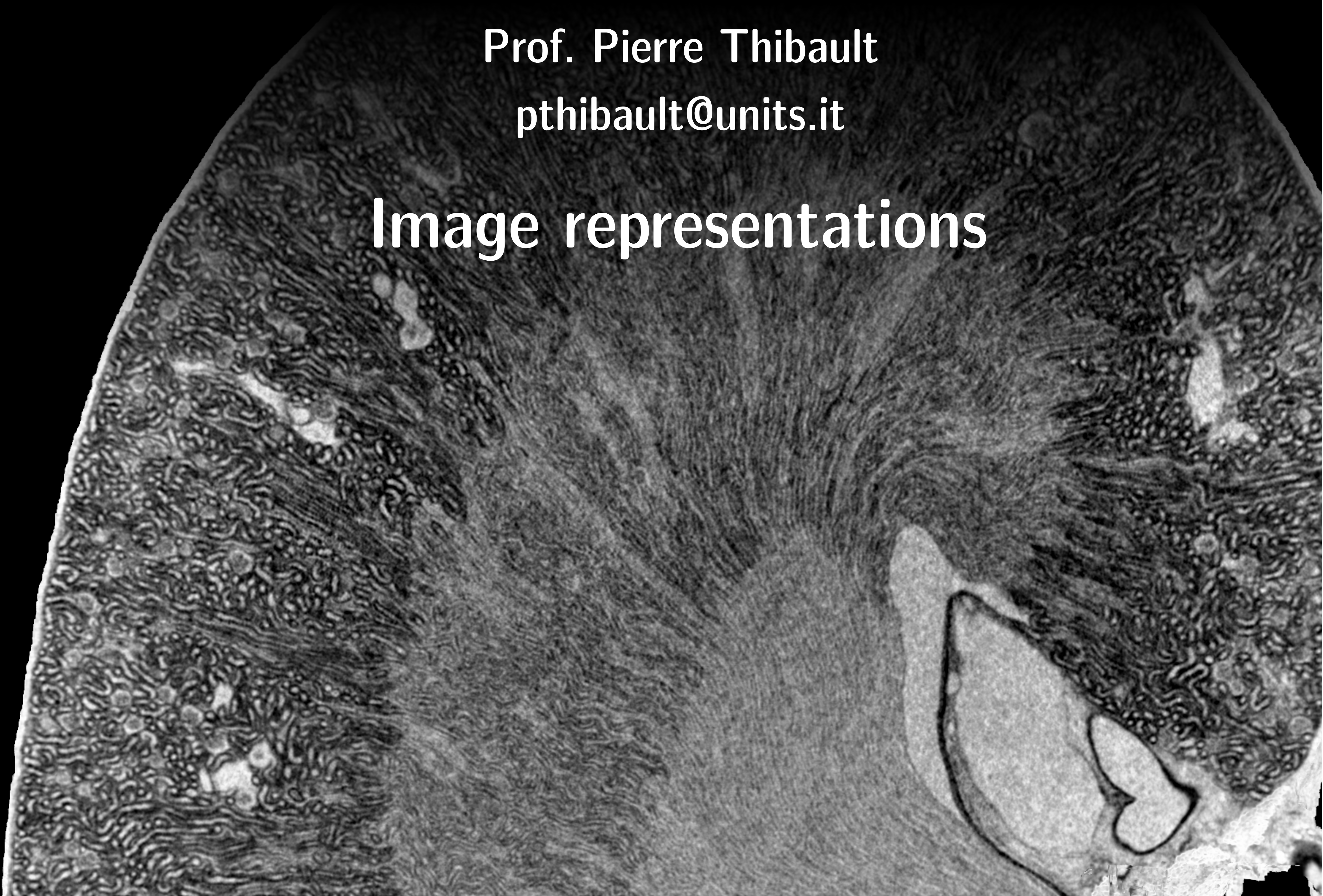
# Image Processing for Physicists

Prof. Pierre Thibault

pthibault@units.it

## Image representations



# Overview

- The Discrete Fourier Transform as a change of basis
- Discrete Cosine Transform
- Windowed Fourier Transform
- Wavelet Transform
- (many others omitted!)

# Image representations

$$f(x,y) = \sum c_n B_n(x,y)$$

$c_n$ : coefficients

$B_n$ : basis function

(most convenient: orthonormal basis)

DFT:

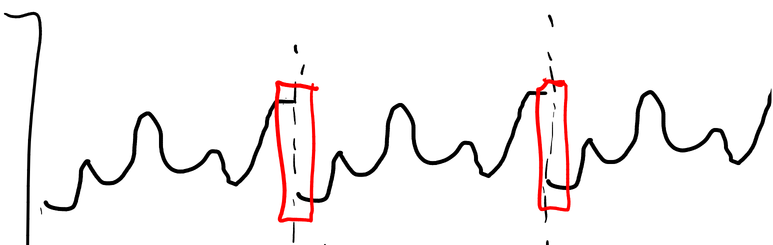
$$f(m,n) = \sum_{k,l} F_{kl} \underbrace{e^{2\pi i \left(\frac{mk}{M} + \frac{nl}{N}\right)}}_{B_{kl}(m,n)\ \text{DFT basis}}$$

$\leftarrow (M,N)$ shaped image

* remember: assumption that $f$ is periodic

1D: $f_n = \sum_k F_k \, e^{2\pi i kn/N}$ $\qquad z = e^{2\pi i/N}$

$$\begin{bmatrix} \\ f \\ \\ \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & 1 & 1 & \cdots & 1 & 1 \\ 1 & z & z^2 & \cdots & & z^{N-1} \\ 1 & z^2 & z^4 & \cdots & & z^{2(N-1)} \\ 1 & \vdots & & & & \\ \vdots & & & & & \\ 1 & & & & & \end{bmatrix}}_{\mathcal{F}^{-1}} \begin{bmatrix} \\ F \\ \\ \end{bmatrix}$$

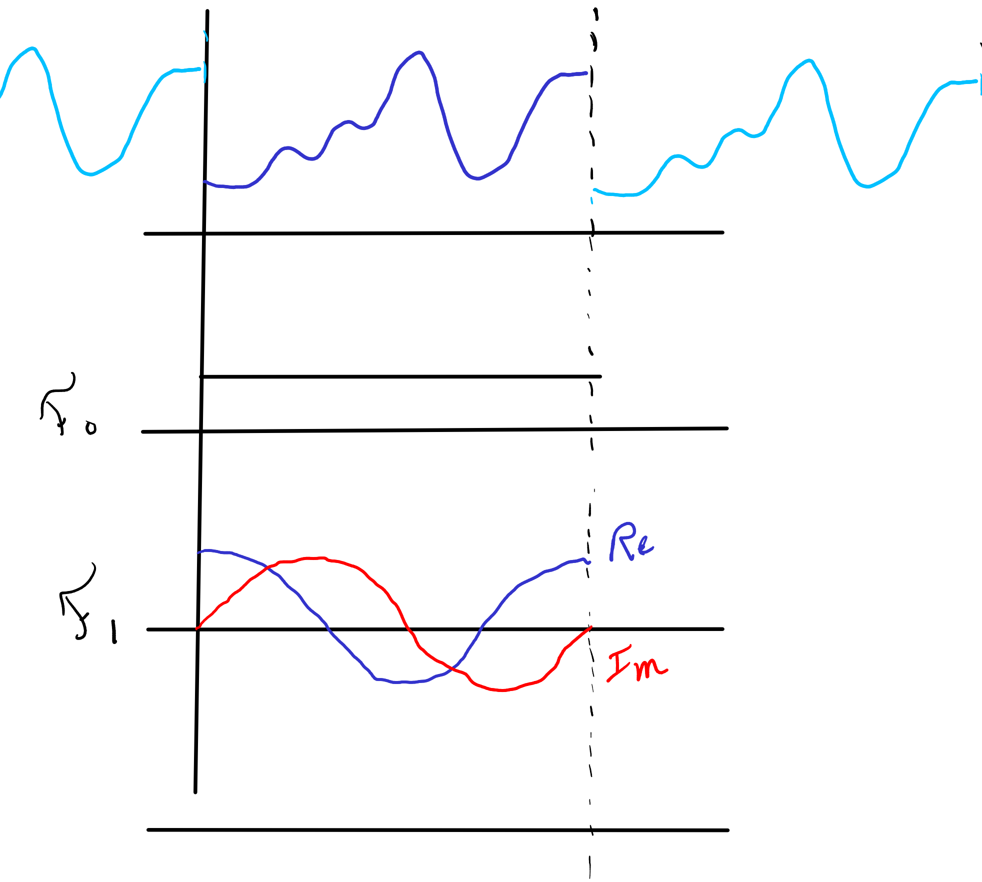


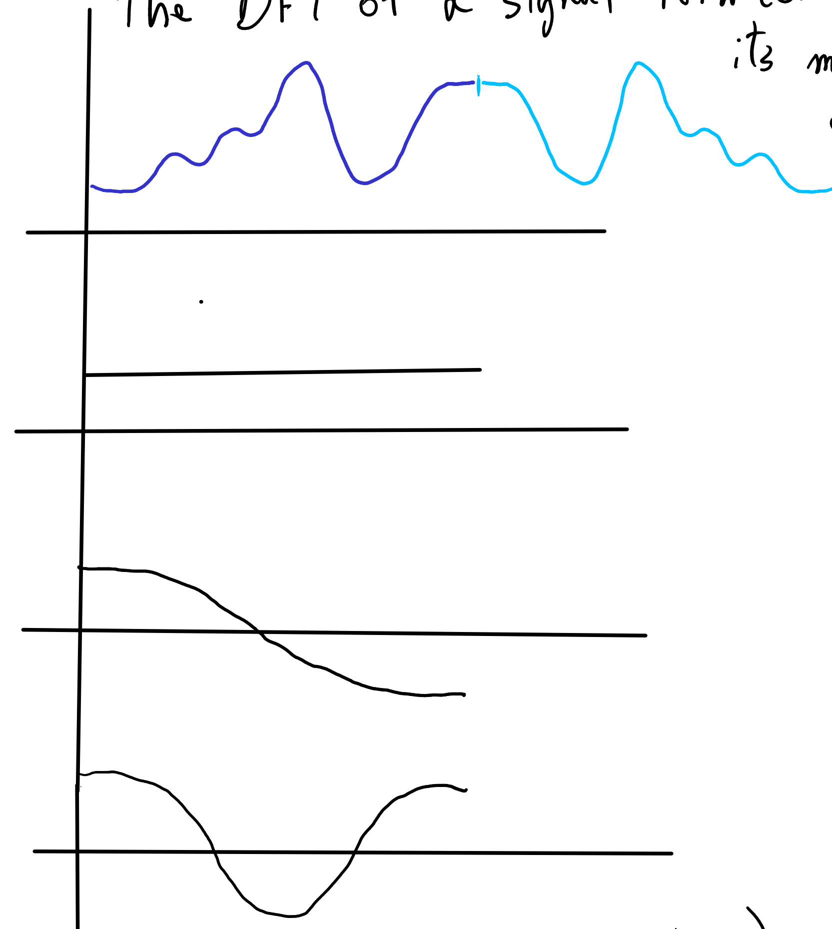
many spatial frequencies are needed to encode a sharp transition (like a step)

# Discrete Cosine Transform

## A variation on the theme of DFT

DCT can be described in terms of the DFT of a signal formed with its mirrored copy



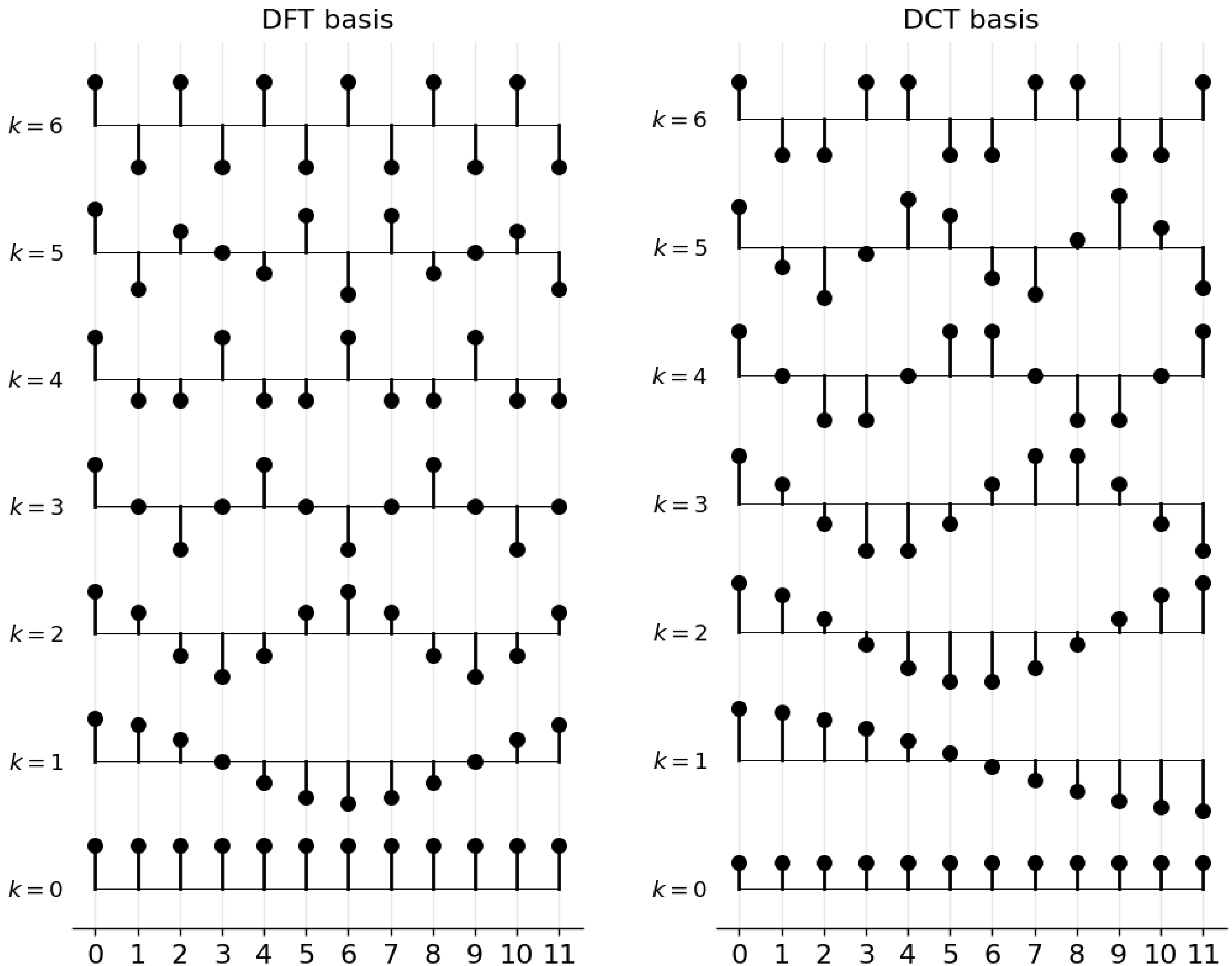



$$\zeta_{kn} = e^{2\pi i k n / N}$$

$$B_k(n) = \cos\left(\left(n+\tfrac{1}{2}\right)\frac{k\pi}{N}\right)$$
$$\left(\times \tfrac{1}{\sqrt{2}} \text{ for } k=0\right)$$

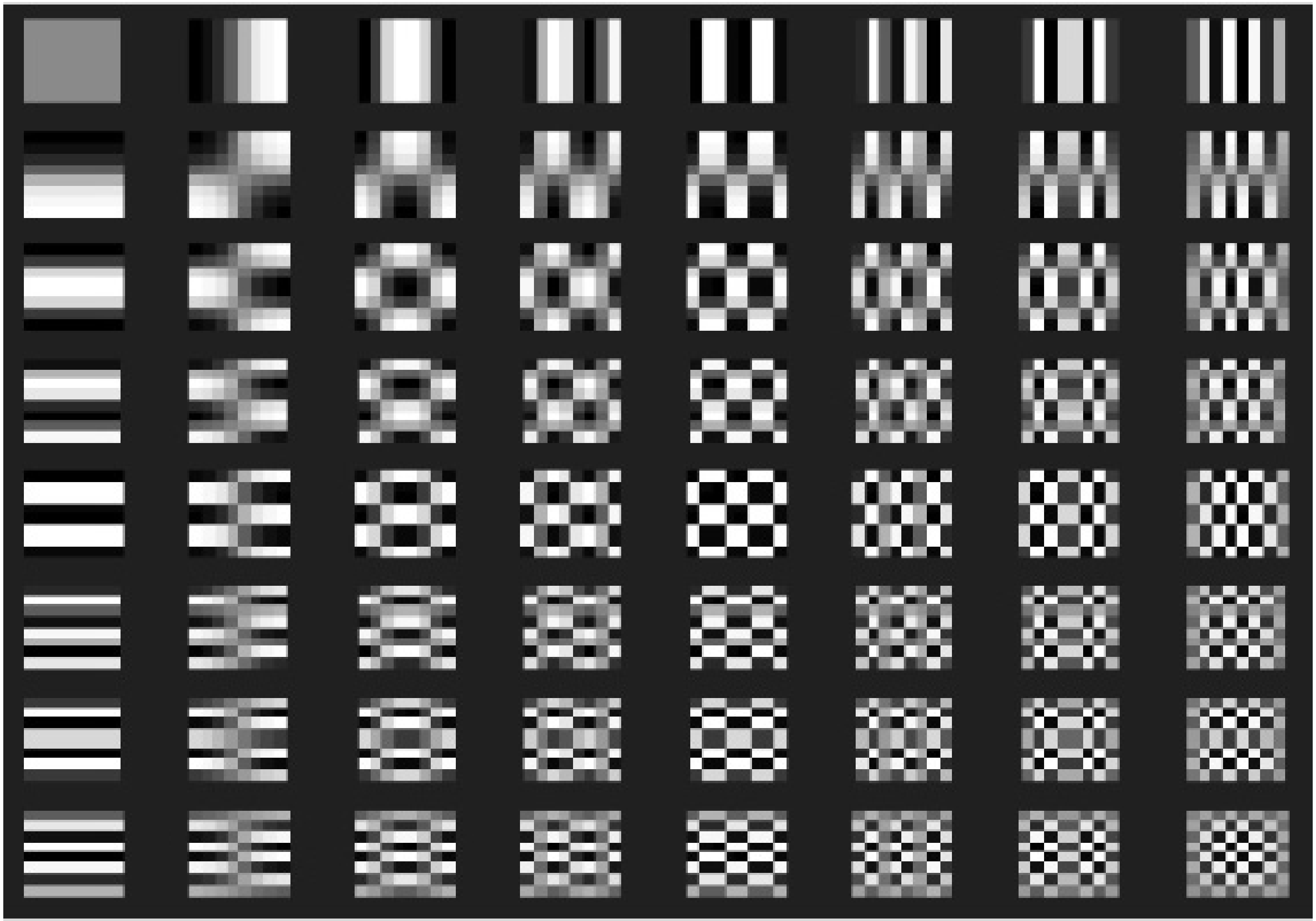
# Discrete Cosine Transform

(N=12)



# Discrete Cosine Transform

64 DCT basis vectors for 8x8 image



# Discrete Cosine Transform

## Image compression

1:1 bit rate

8:1 bit rate

JPEG compression

↖ keep in average 8 most significant coefficients

32:1 bit rate

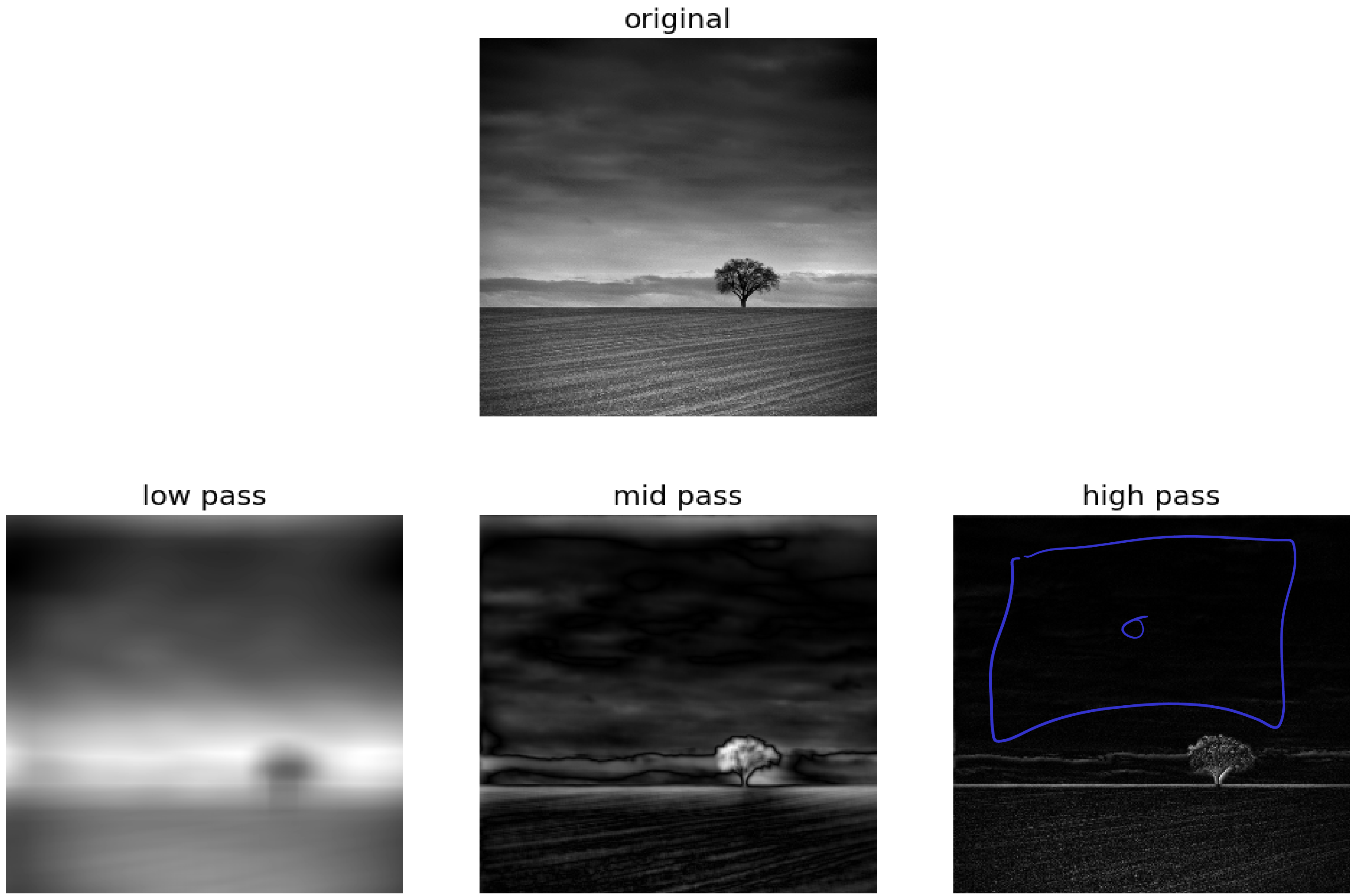
128:1 bit rate

lossy compression
information is discarded

# Historical overview

- 1822 Fourier: Fourier transform
- 1946 Gabor: “Gabor transform”, Short-time Fourier transform (STFT)
- 1974 Ahmed, Natarajan & Rao: Discrete Cosine Transform
- 1980s Morlet, Mallat, Daubechies, … : Wavelets

# Bandpass filtering

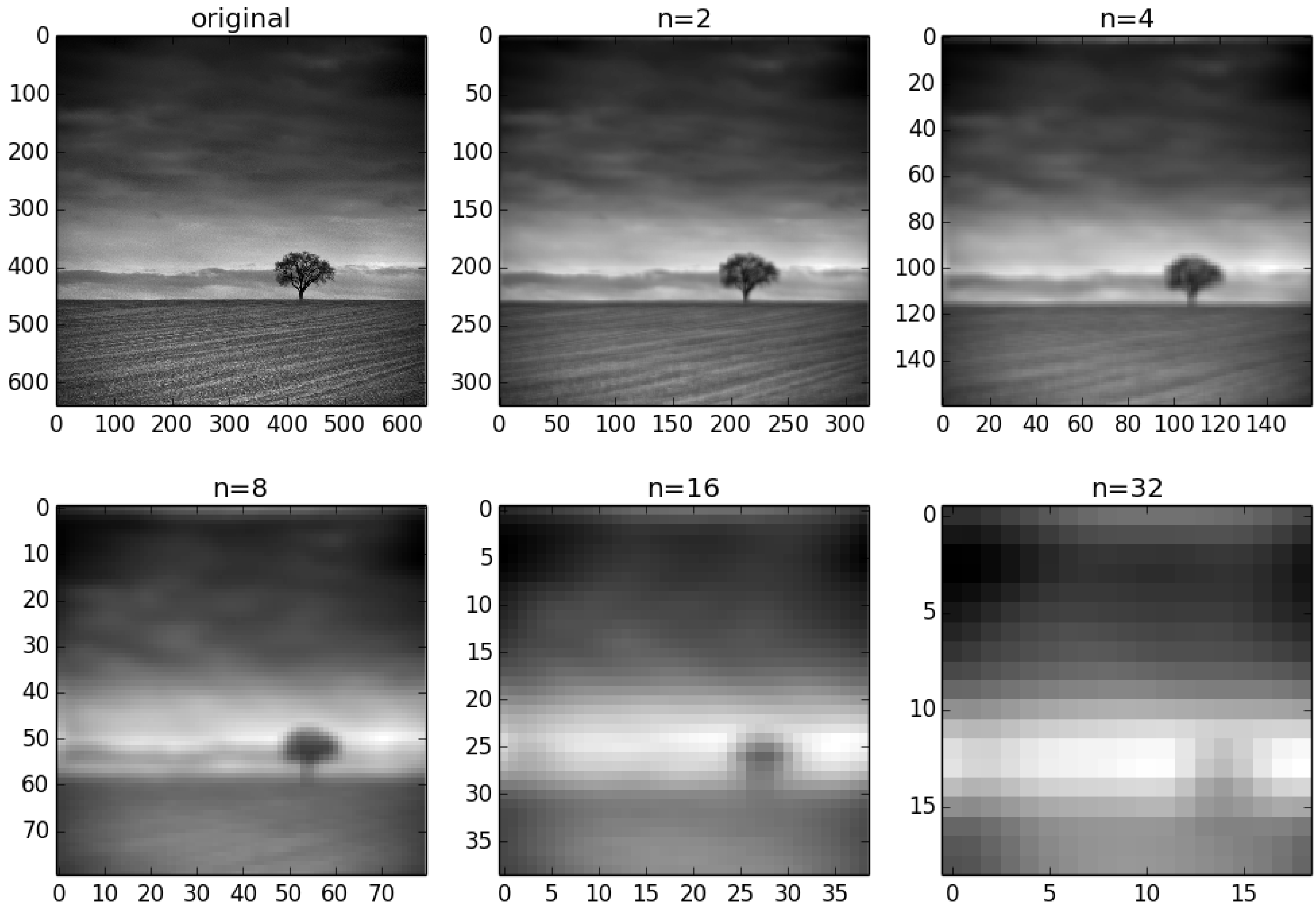



Don't need high spatial resolution

Need high spatial resolution

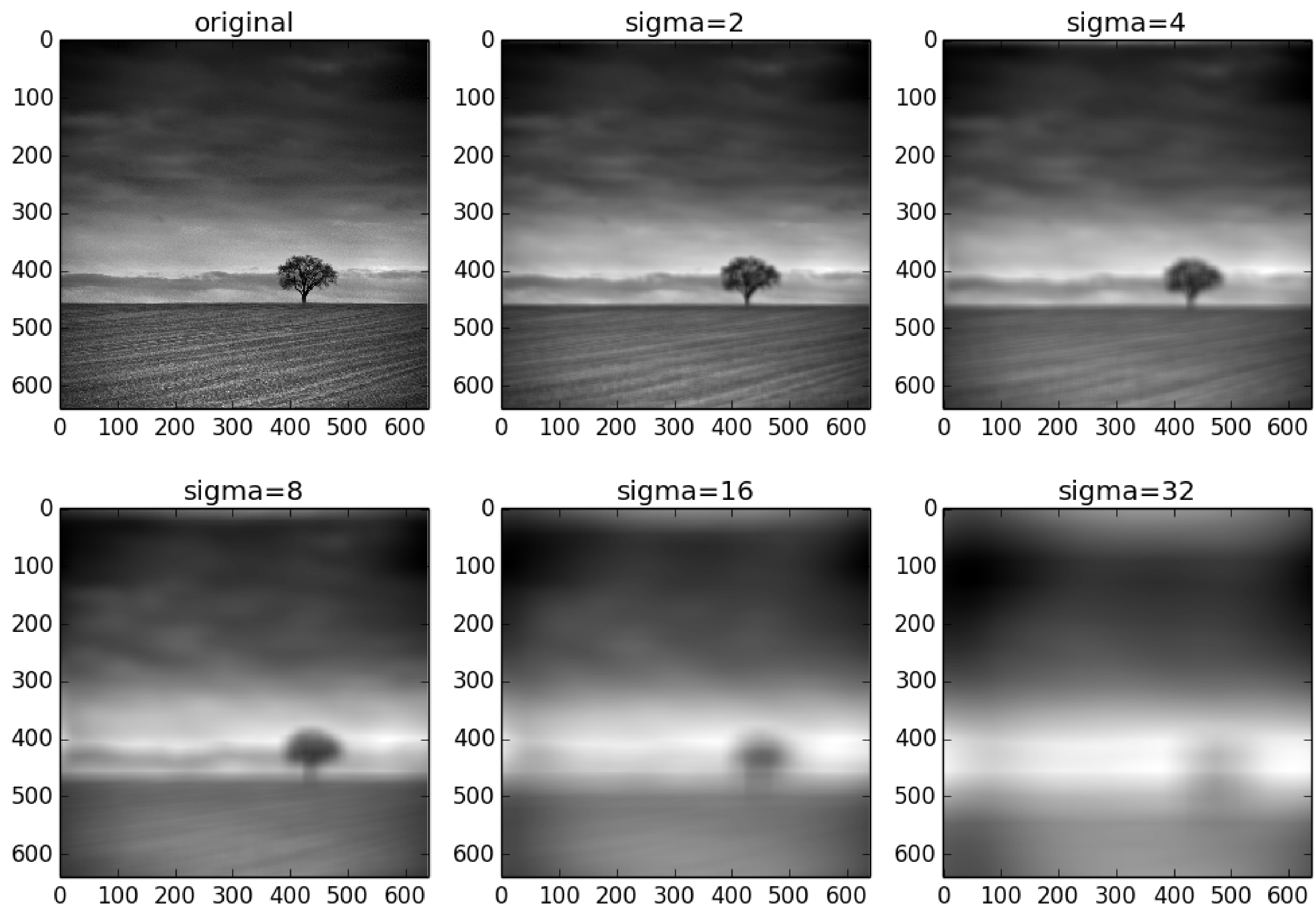
# Multiresolution analysis

Subsampling (taking every n$^{th}$ pixel) successively reduces high frequency content



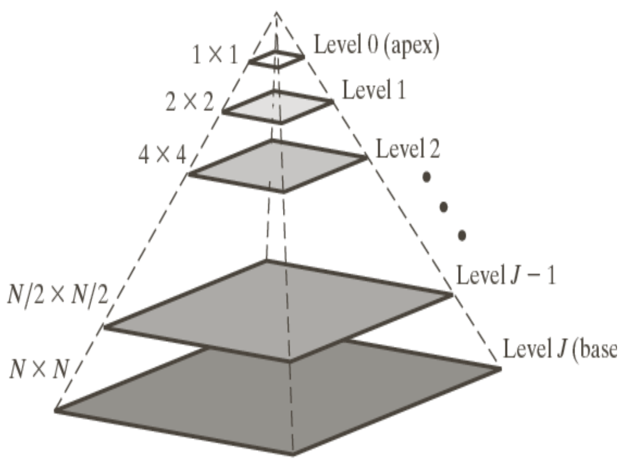
# Multiresolution analysis

Multiple filtering with Gaussian filters, sigma determines resolution



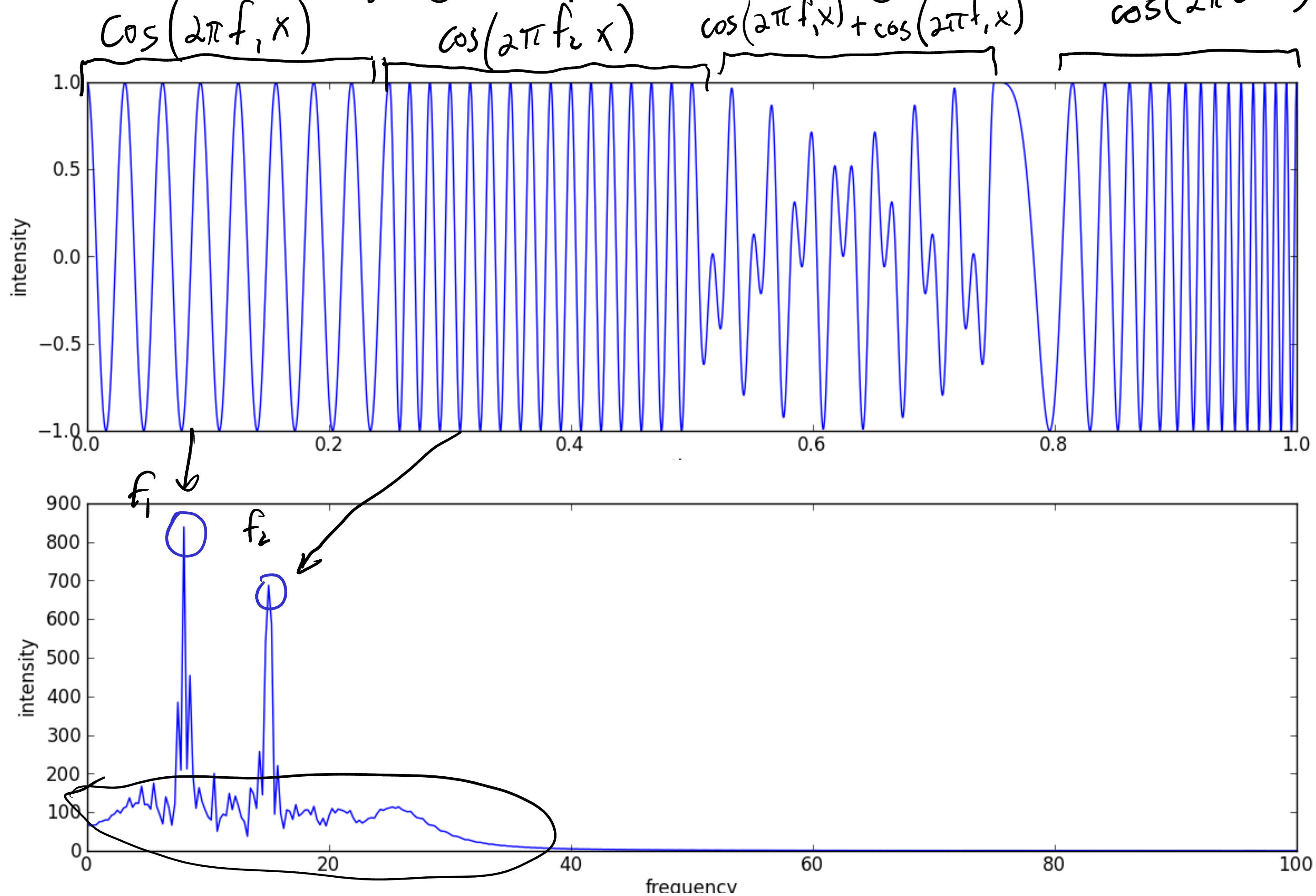
# Pyramid representation

Scale-space representation, pyramidal representation




# Stationary vs. non-stationary signals

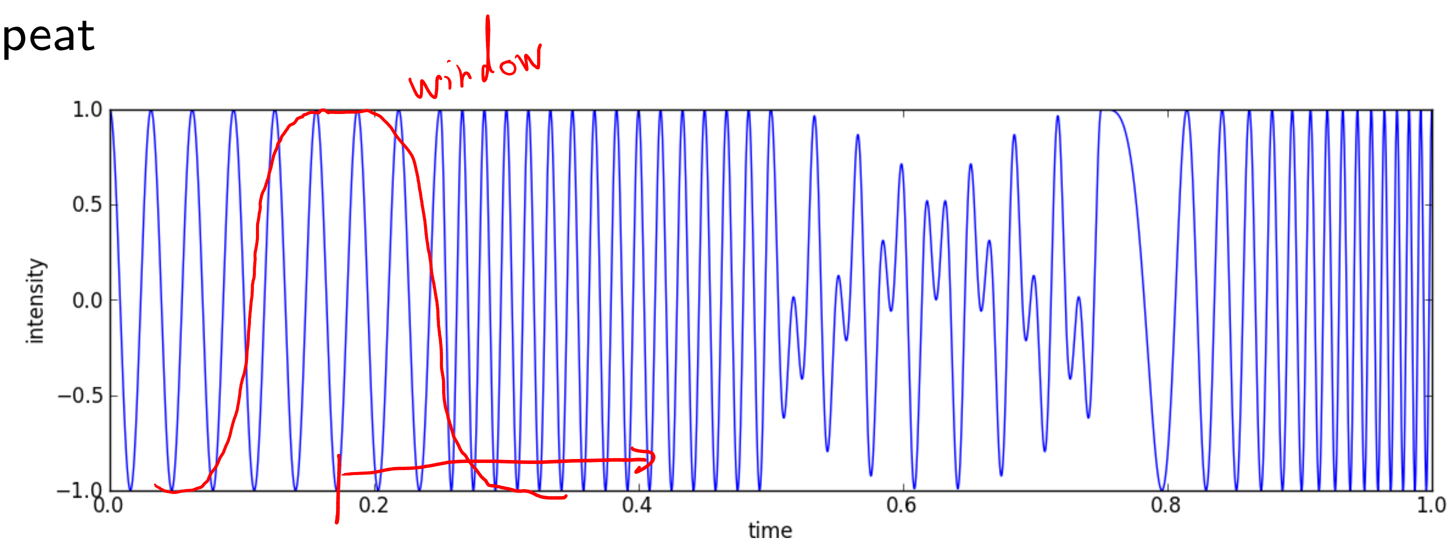
- Stationary signals: frequency doesn't change over time (spatially over the image)
- Non-stationary signals: frequency changes over time (spatially over the image)
- Examples of non-stationary signals: speech, most images




FT insufficient to localize the frequencies in our signal (image)

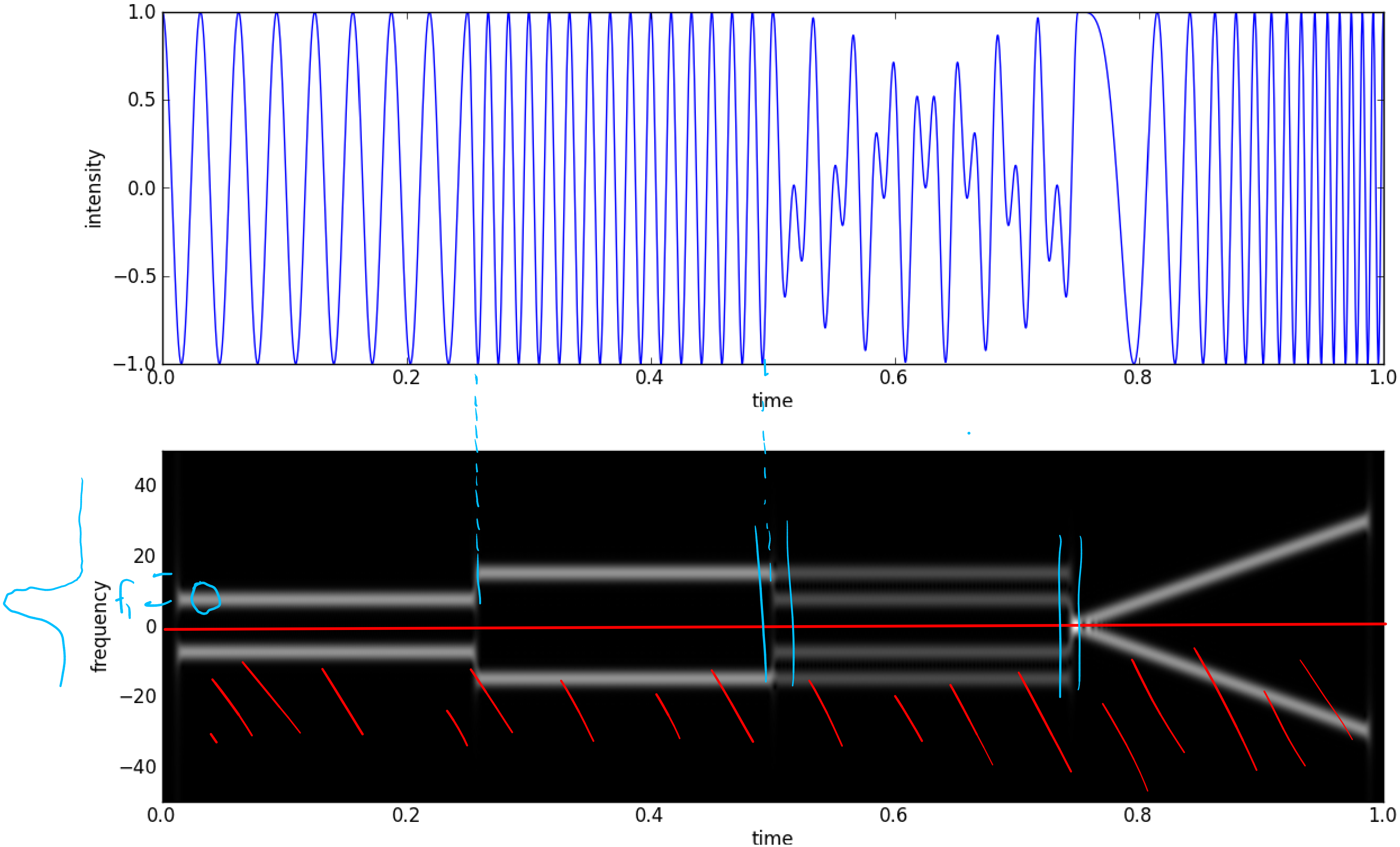
# Windowed Fourier transform

- Windowed Fourier transform is part of the field of "time-frequency analysis"
- Also known as Short-time Fourier Transform (STFT)
- Time-frequency representations are used in many different contexts (Audio, image processing/optics, quantum mechanics)
- Idea: slice up signal into small parts, analyze each separately
  - Multiply with window function w (of width d) at position x0
  - Take Fourier transform of result
  - Slide window to new position
  - repeat



# Analogy to audio signals

# Spectrogram

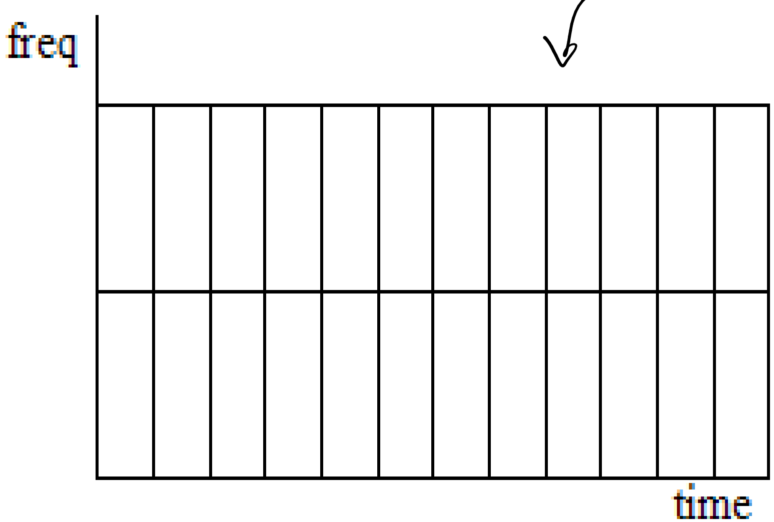


# Uncertainty relation

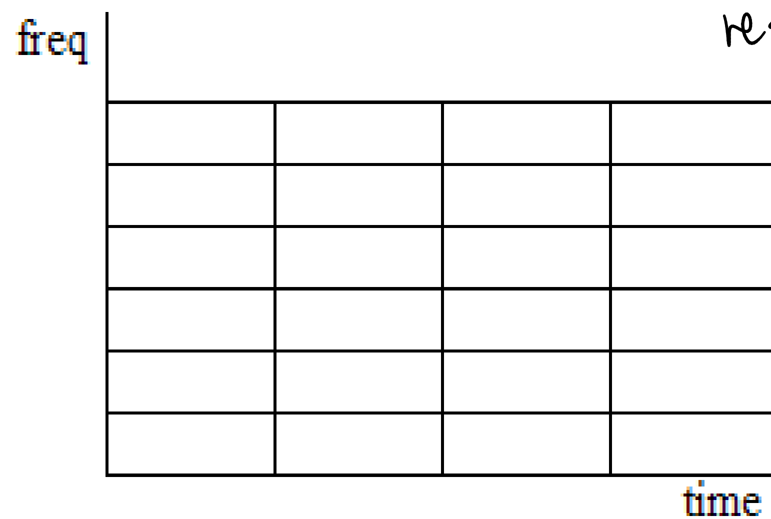
$$\sigma_s \, \sigma_f \geq \frac{1}{4\pi}$$

- Finite area in the time-frequency plane

better time resolution




better frequency resolution




- This is limitation of WFT and hence development of **wavelets**

# Continuous wavelet transform (WT)

"mother" wavelet

$$\psi_{s,x_0} = \frac{1}{\sqrt{s}} \psi\left(\frac{x - x_0}{s}\right)$$

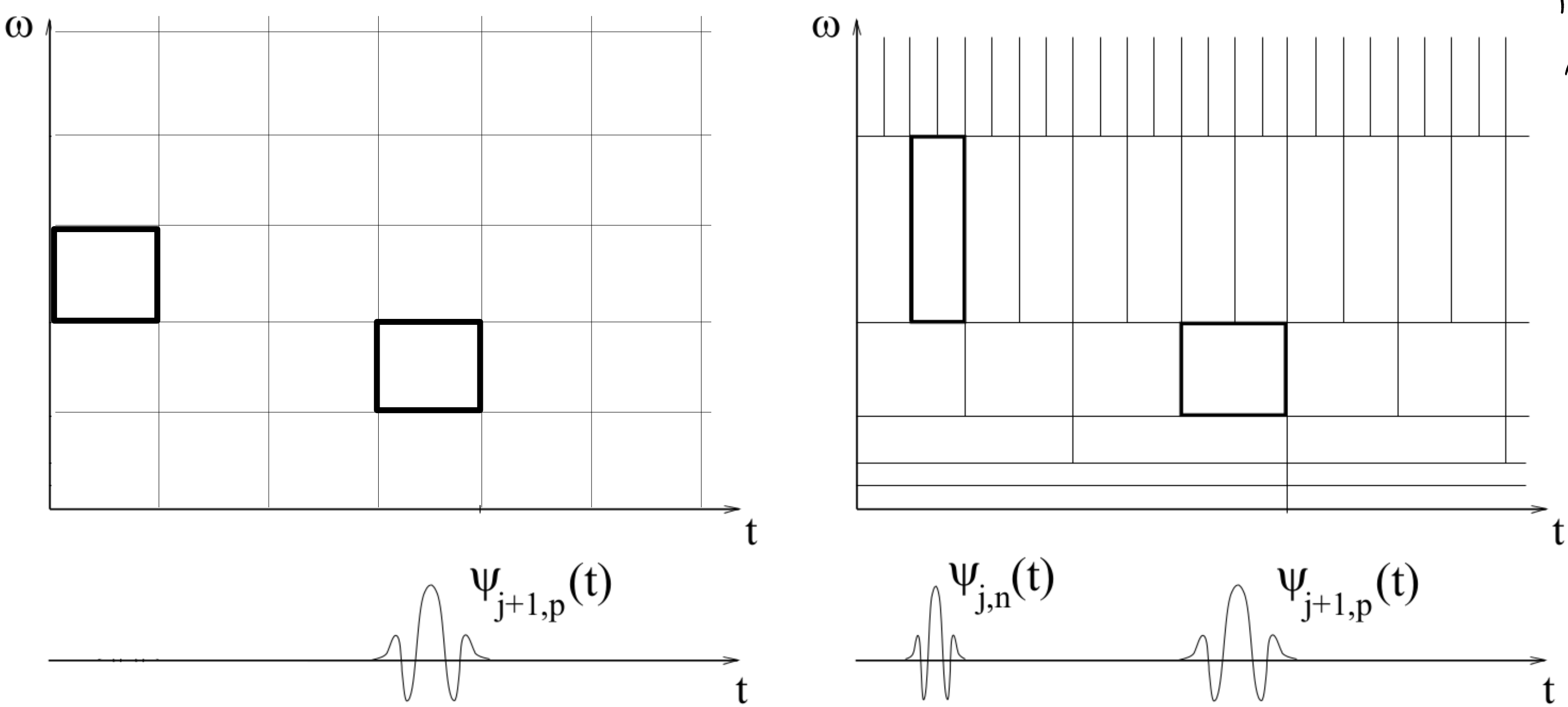
- Parameters: translation and scaling

$$WT_{s,x_0}\{f\} = \int_{-\infty}^{\infty} f(x)\, \psi_{s,x_0}(x)\, dx$$

← no $e^{2\pi i x}$ anymore (not F.T.)

- Analyze signal at <u>different scales</u> instead of <u>different frequencies</u>

oscillation behavior is in the mother wavelet

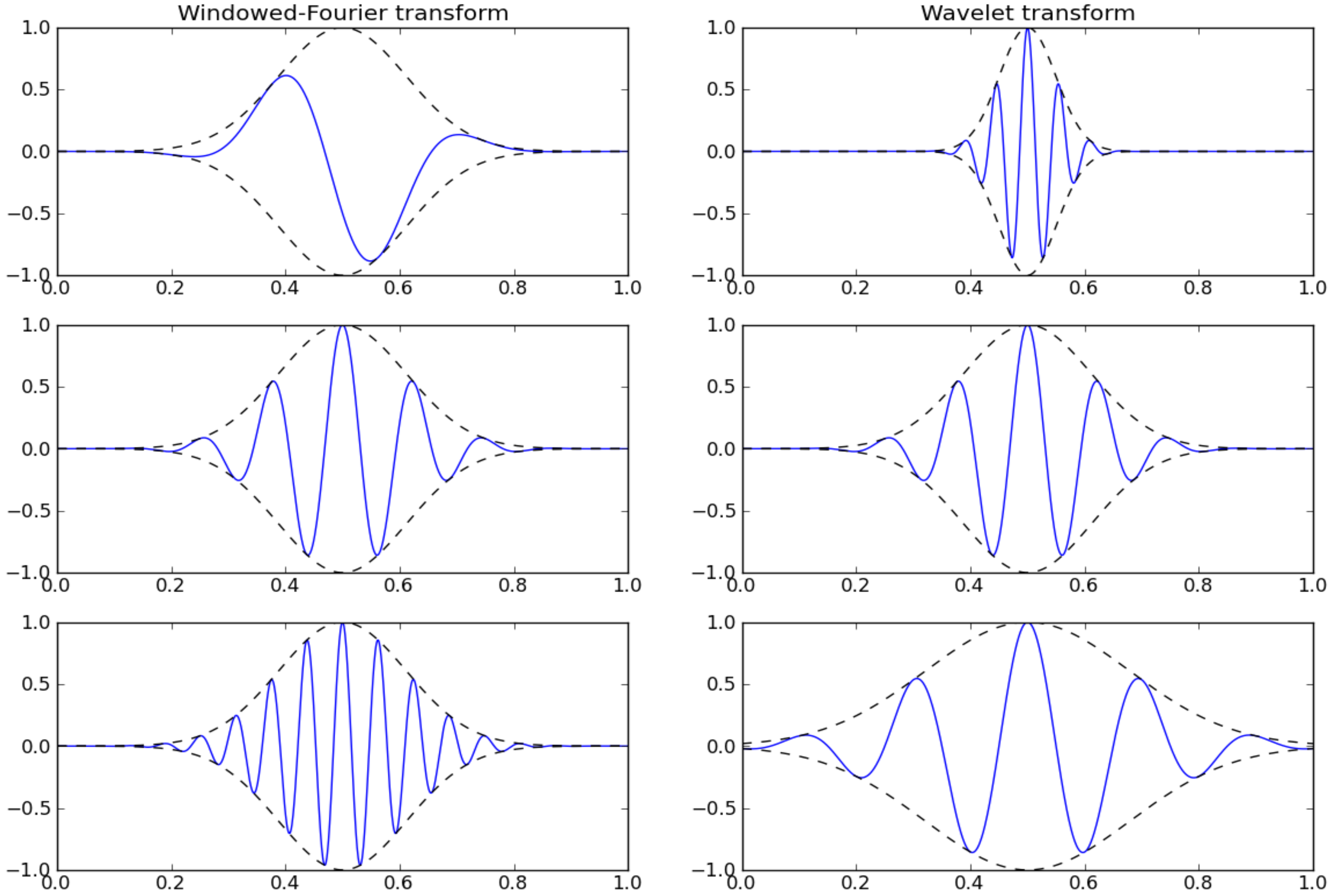



Source: Mallat, "A wavelet tour of signal processing"

# WFT vs WT

WFT - keep window width constant
- change modulation

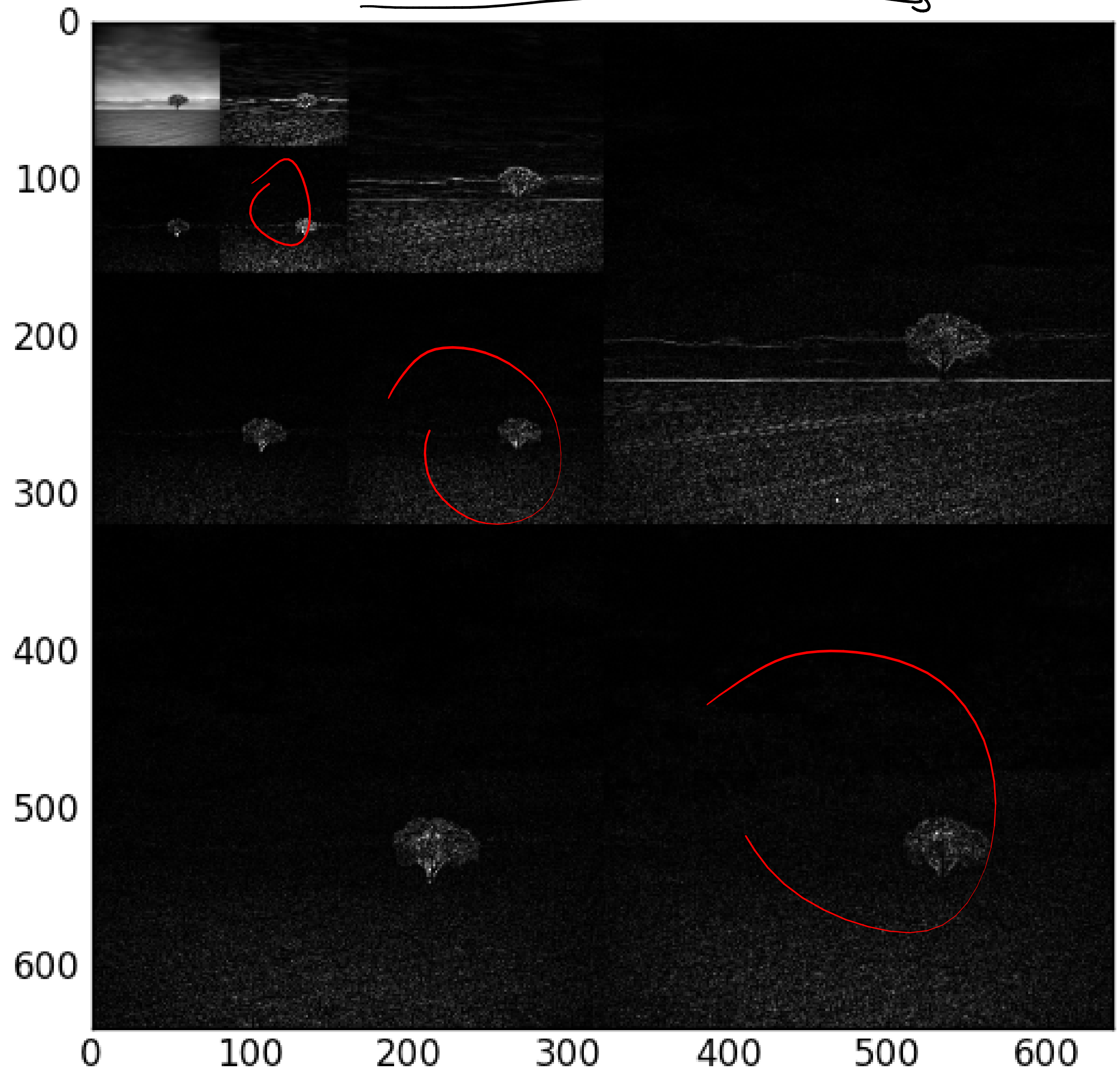
Wavelet - keep shape constant
- change scale



# Discrete Wavelet decomposition of image

- Perform each DWT, collect and tile all coefficients
- Here: 3 level decomposition

used for denoising or as a "regularizer" ← to be defined later




vertical variations

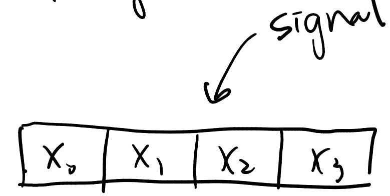
horizontal

"diagonal"

JPEG 2000 is based on wavelets

# Haar wavelet

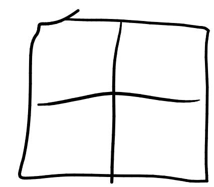
1D $N = 4$ signal

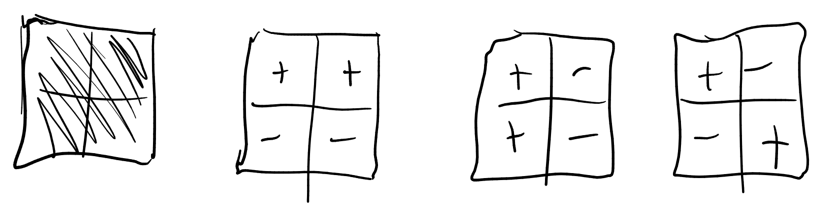



$$H_4 = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 \end{bmatrix}, \begin{bmatrix} X \end{bmatrix} = \begin{bmatrix} x_0 + x_1 + x_2 + x_3 \\ x_0 + x_1 - (x_2 + x_3) \\ x_1 - x_1 \\ x_2 - x_3 \end{bmatrix}$$

$$H_4 = \frac{1}{2} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ \sqrt{2} & -\sqrt{2} & 0 & 0 \\ 0 & 0 & \sqrt{2} & -\sqrt{2} \end{bmatrix}$$

2D $2 \times 2$ signal





# Summary

- Images can be represented by different basis functions.
- Fourier basis: localized in frequency, delocalized in real space.
- Windowed Fourier Transform: localized – to some extent – in both spaces
- Wavelet analysis decomposes a signal in position and scale (instead of position and frequency as for WFT).
- Sparse representations are representations in which the image content is represented by a few relevant coefficients, while the other pixels are close to zero
- Sparse representations have advantages for compression, denoising, ...