M2RI, Parcours Images et Données

Module Acquisition et Représentation des Données

Traitements et Transformations

Rémi Gribonval

Projet PANAMA

IRISA

Bureau C319

remi.gribonval@inria.fr

http://people.irisa.fr/Remi.Gribonval

IRISA

UNE UNITÉ DE RECHERCHE À LA POINTE DES SCIENCES ET DES TECHNOLOGIES DE L'INFORMATION ET DE LA COMMUNICATION

Quatrième partie

UNIVERSITÉ DE RENNES 1 · INSA · CNRS · INRIA RENNES

Summary of last class

- **Problem**: represent the time-varying frequency content of time series
- **Tool**: Short Time Fourier Transform (STFT)
  - definition
  - role of the analysis window (size & shape)
  - perfect reconstruction with synthesis window
- **Applications**: time-frequency processing
  - Effect of redundancy of STFT

Today's program

- **Solutions** of Exercise
- **From signals (1D) to images (2D)**
- **Tool**: multiresolution analysis (1D & 2D)
  - 2D Fourier transform
  - pyramidal representations & wavelets
- **Applications**: Denoising & Inpainting
- **Concept of sparsity**

Solutions of Exercise (STFT)

Example : Hann window

- Analysis window, synthesis window

$$w[n] = \frac{1}{2}\left(1 - \cos 2\pi \frac{n}{N}\right), \; 0 \leq n < N \qquad w'[n] = 1$$

- Half-overlapping window, hop size = $T = N/2$
- **Exercice 1**: check the reconstruction condition
  - Hint : $(1 - \cos t)/2 = \sin^2 t/2$
- **Exercice 2**: idem with sine window

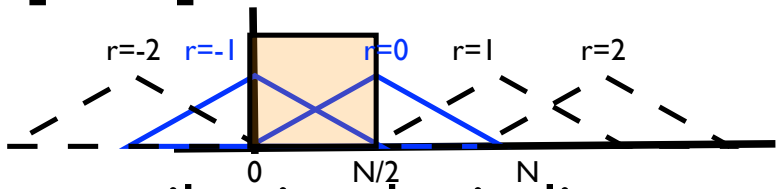
$$w[n] = w'[n] = \sin \pi \frac{n}{N}, \; 0 \leq n < N$$

Solution

- What we need to check

$$f[n] := \sum_{r=-\infty}^{+\infty} w[n-rT]w'[n-rT] = 1, \forall n \in \mathbb{Z}$$

- f[n] is T-periodic:so only need to check $f[n] = 1, 0 \leq n < T = N/2$
- Support of w[n-rT] = $[\![0, N-1]\!]$



- Sum = only contribution by indices r=-1,0

$$\begin{aligned} w[n+T] + w[n] &= \sin^2(\pi(n+T)/N) + \sin^2(\pi n/N) \\ &= \sin^2(\pi n/N + \pi/2) + \sin^2(\pi n/N) \\ &= \cos^2(\pi n/N) + \sin^2(\pi n/N) \end{aligned}$$

From signals to images

2D Fourier transform

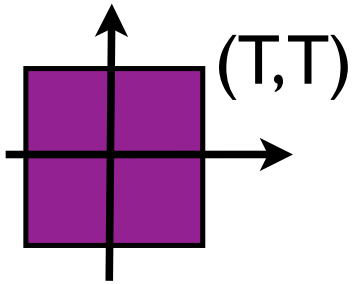
- Definition $X(\vec{f}) := \int\int x(\vec{p}) e^{-2i\pi f_1 p_1 - 2i\pi f_2 p_2} dp_1 dp_2$
- Same properties as 1D Fourier transform
- Rotation $R_\theta := \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$

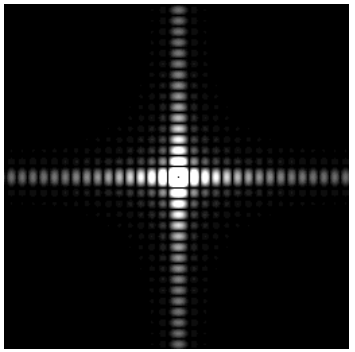
  $x(R_{-\theta}\vec{p})$ $\qquad$ $X(R_{-\theta}\vec{f})$
- Separable images / tensor products

  $x_1(p_1) \cdot x_2(p_2)$ $\qquad$ $X_1(f_1) \cdot X_2(f_2)$

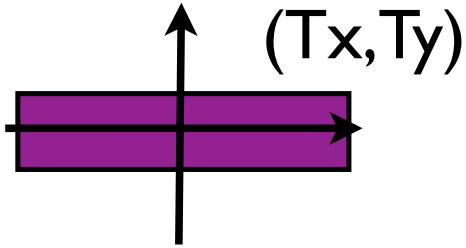
Example [exercice]

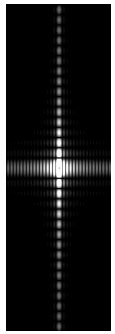
- Square



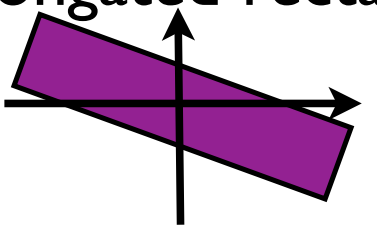


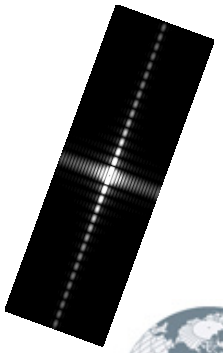
- Rectangle





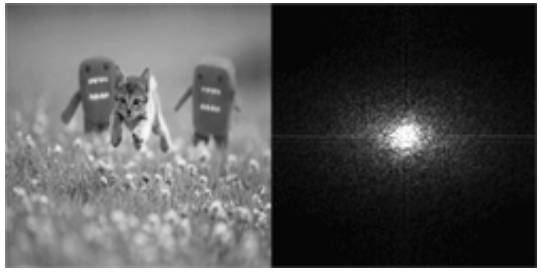
- Rotated elongated rectangle





Example

- A natural image and its Fourier transform



  - Low frequency content is dominant
  - Some horizontal & vertical edges
  - Spatial information is hard to extract

Multiresolution representations

Multiscale image representations

- Low frequency content = low-resolution image approximation = coarse image
- High-frequency content = details
- Idea = coarse to fine description of images
  - pyramidal representation [Burt & Adelson]
  - wavelets [Mallat, Daubechies, Meyer, Vetterli, ...]

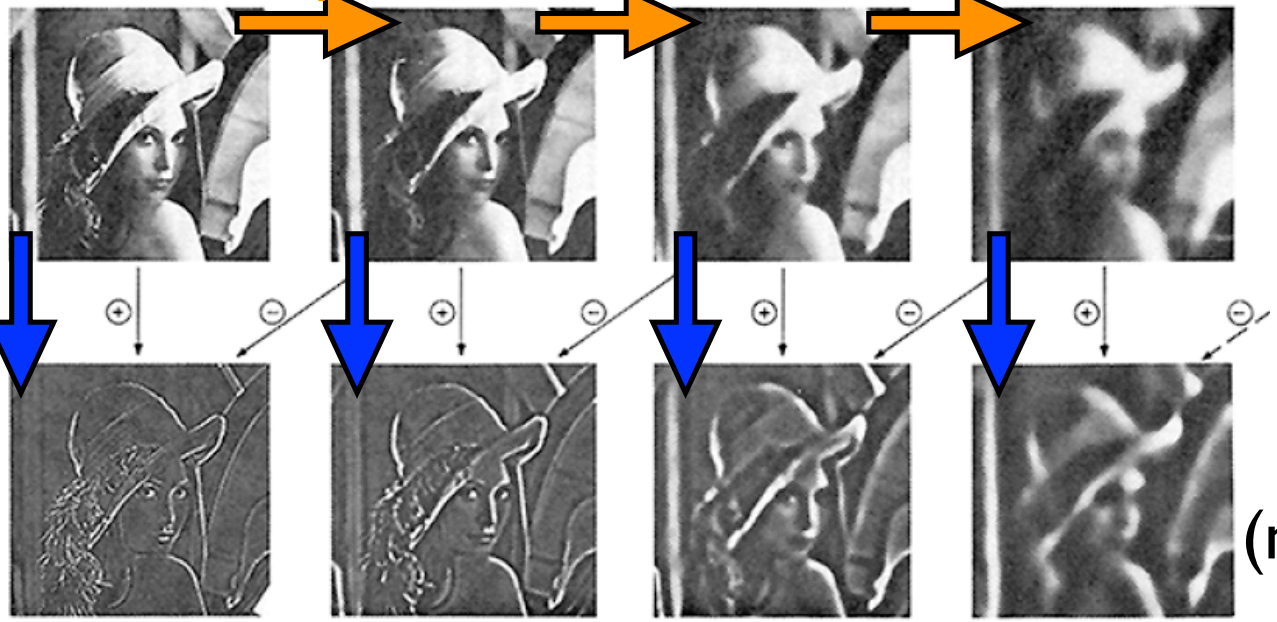
Pyramidal representation

Small scale
High resolution

Large scale
Low resolution

Low pass

High pass



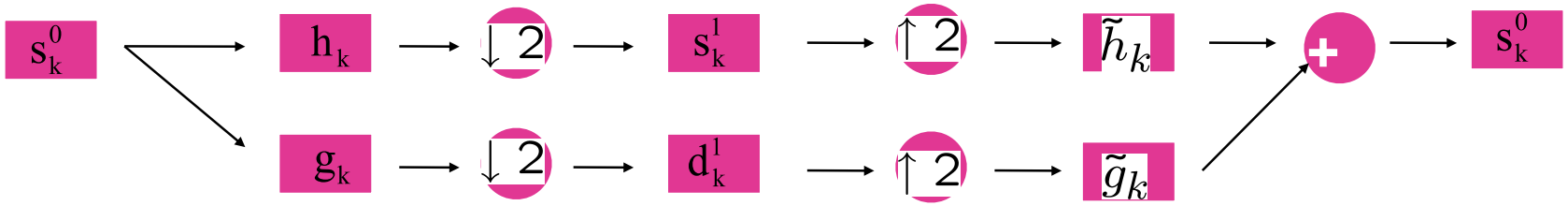
Coarse approx.

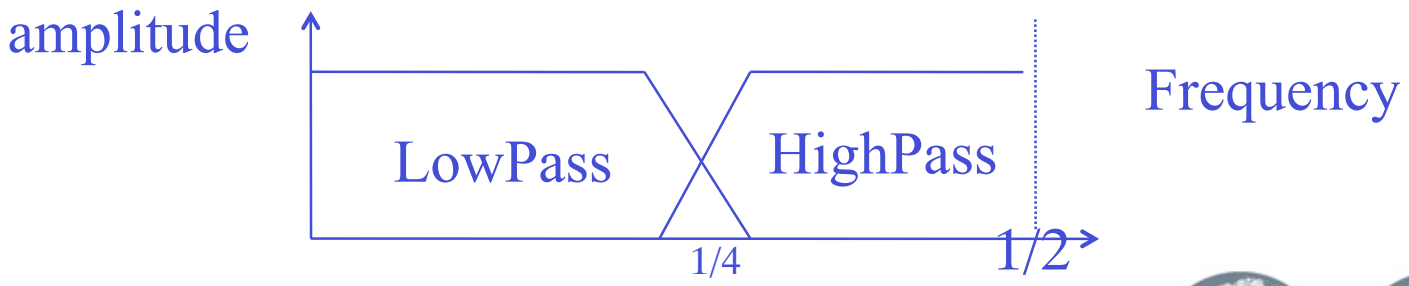
Details (many edges)

1D wavelets

1D subsampled filterbanks

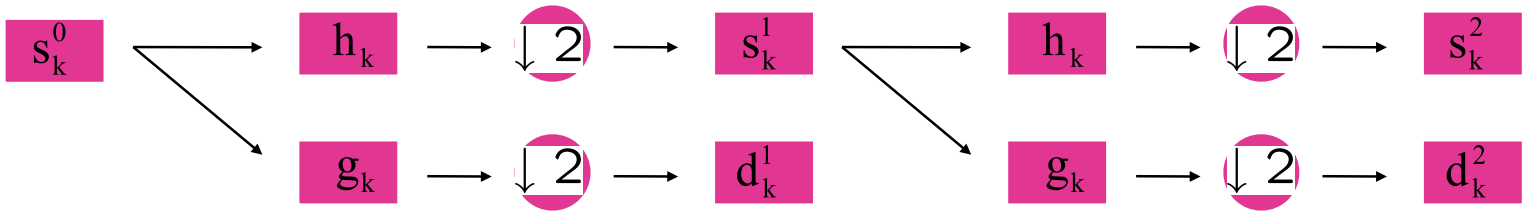
- Subsampling at each low-pass / high-pass to remove unnecessary redundancy





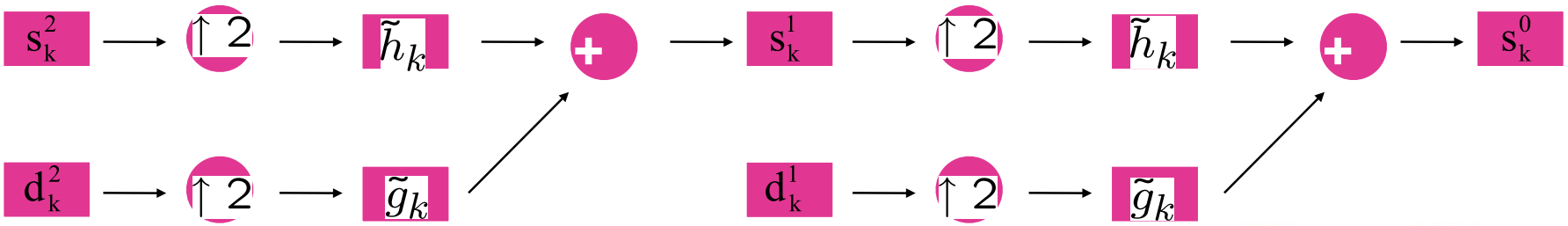
1D Multiscale decomposition

- Fast Analysis



Wavelet coefficients at scale 1

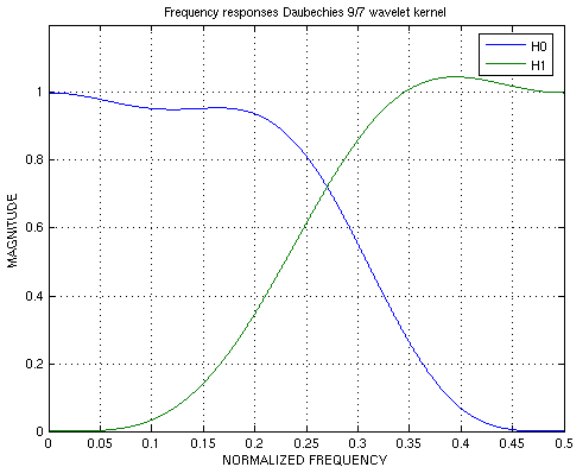
- Fast Synthesis



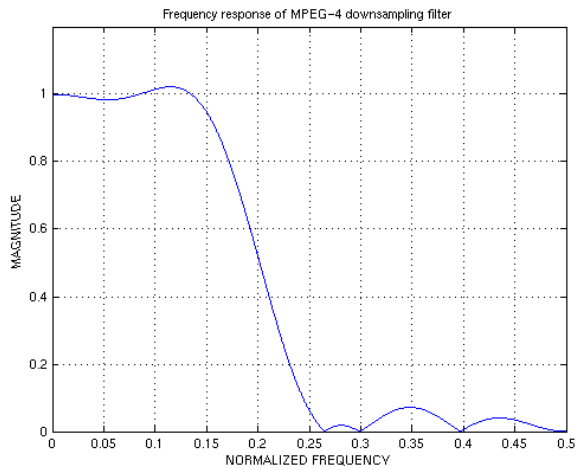
1D wavelet filters

- To get perfect reconstruction : Quadrature Mirror Filters (QMF)
- Examples :

Daubechies 9/7



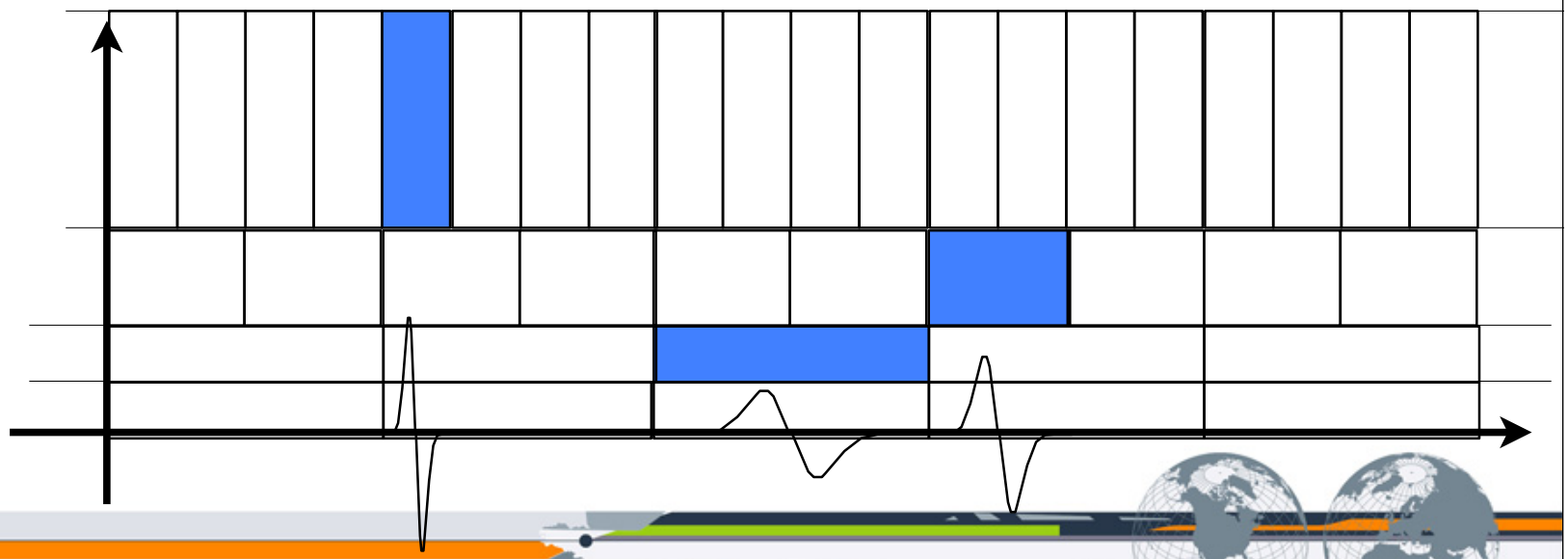
MPEG-4 downsampling (N3908)



Time-frequency tiling

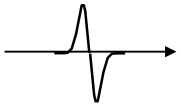
$$2^{j/2}\psi(2^j t - k)$$

- Dilation + translation of mother wavelet $\psi(t)$
- Time-frequency localization of wavelets



1D orthonormal wavelets

- Detail coefficients at scale j $d_{j,k} = \langle x, \psi_{j,k} \rangle$

  $\psi_{j,k}(t) = 2^{j/2}\psi(2^j t - k)$ mother wavelet function 

- Coarse coefficients at scale j $s_{j,k} = \langle x, \phi_{j,k} \rangle$

  $\phi_{j,k}(t) = 2^{j/2}\phi(2^j t - k)$ scaling function 

- Reconstruction $x = \sum_{j=0}^{J} \sum_{k} \langle x, \psi_{j,k} \rangle \psi_{j,k} + \sum_{k} \langle x, \phi_{J,k} \rangle \phi_{J,k}$
- Orthonormal Basis $\{\psi_{j,k}\}_{j,k} \cup \{\phi_{J,k}\}_{k}$

Choice of the mother wavelet

- Through the choice of H and G QMF filters
- Main characteristics to consider
  - vanishing moments = level of oscillations

$$\int \psi(t) t^n dt = 0, 0 \leq n \leq N \Longrightarrow \langle x, \psi_{j,k} \rangle \approx 0$$

on smooth parts of signal/.image

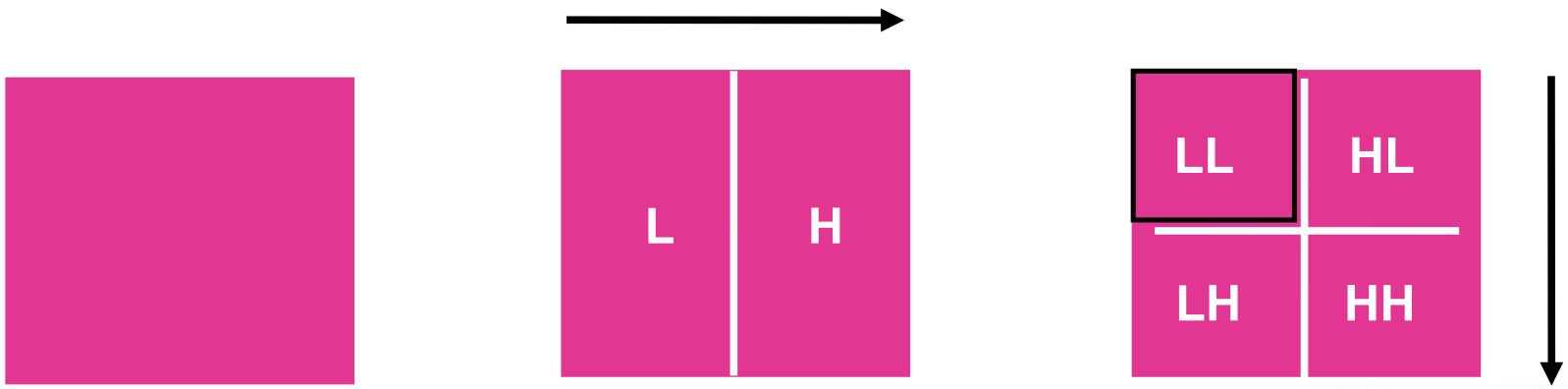
  - length [shorter = faster convolution]
  - regularity

2D wavelets

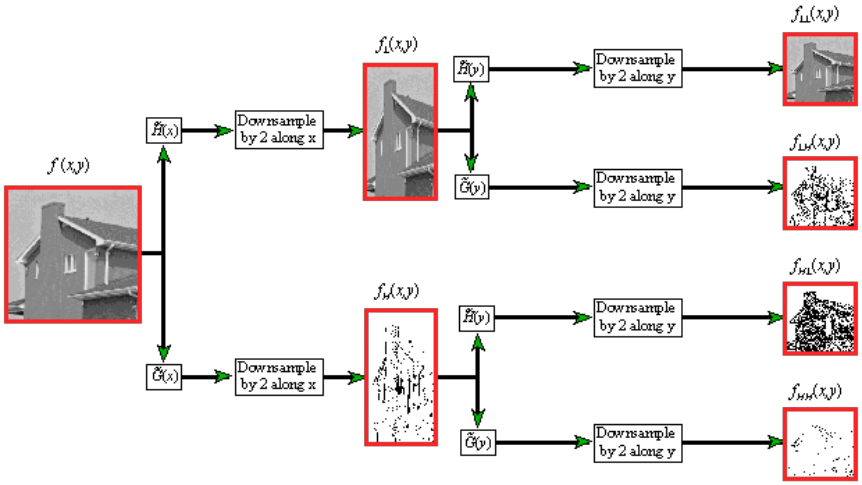
2D orthonormal wavelets

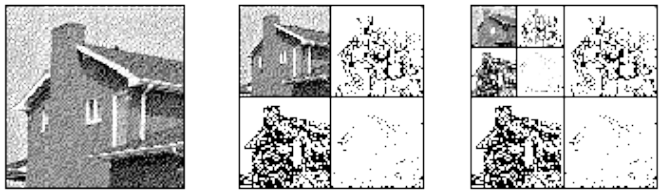
- 1D filters applied alternatively on rows/ columns of the image, considered as 1D signals

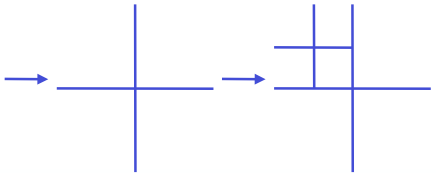
H=**L**ow-pass/G=**H**igh-pass



Example

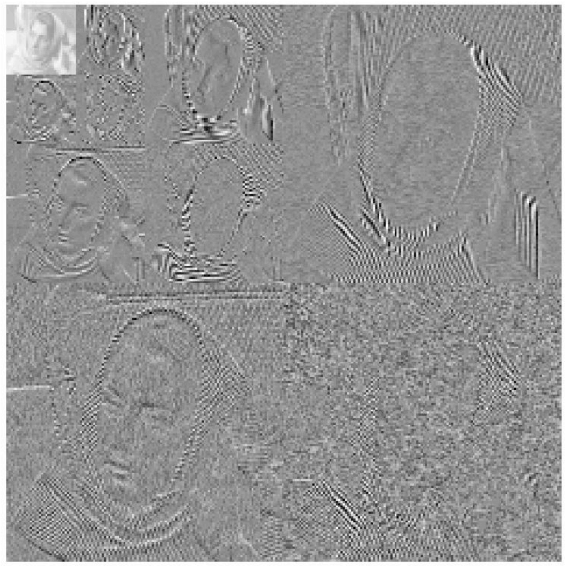

f (x,y)
H̃(x)
G̃(x)
Downsample by 2 along x
Downsample by 2 along x
f_L(x,y)
f_H(x,y)
H̃(y)
G̃(y)
H̃(y)
G̃(y)
Downsample by 2 along y
Downsample by 2 along y
Downsample by 2 along y
Downsample by 2 along y
f_LL(x,y)
f_LH(x,y)
f_HL(x,y)
f_HH(x,y)

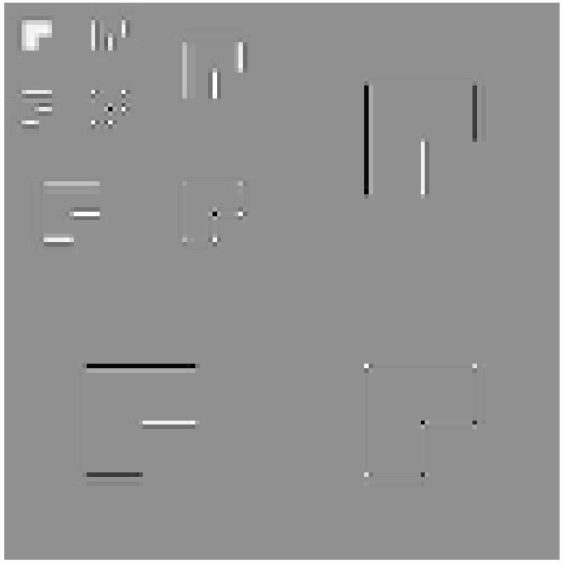




Example of 2D wavelet transform

- 3 layers of decomposition

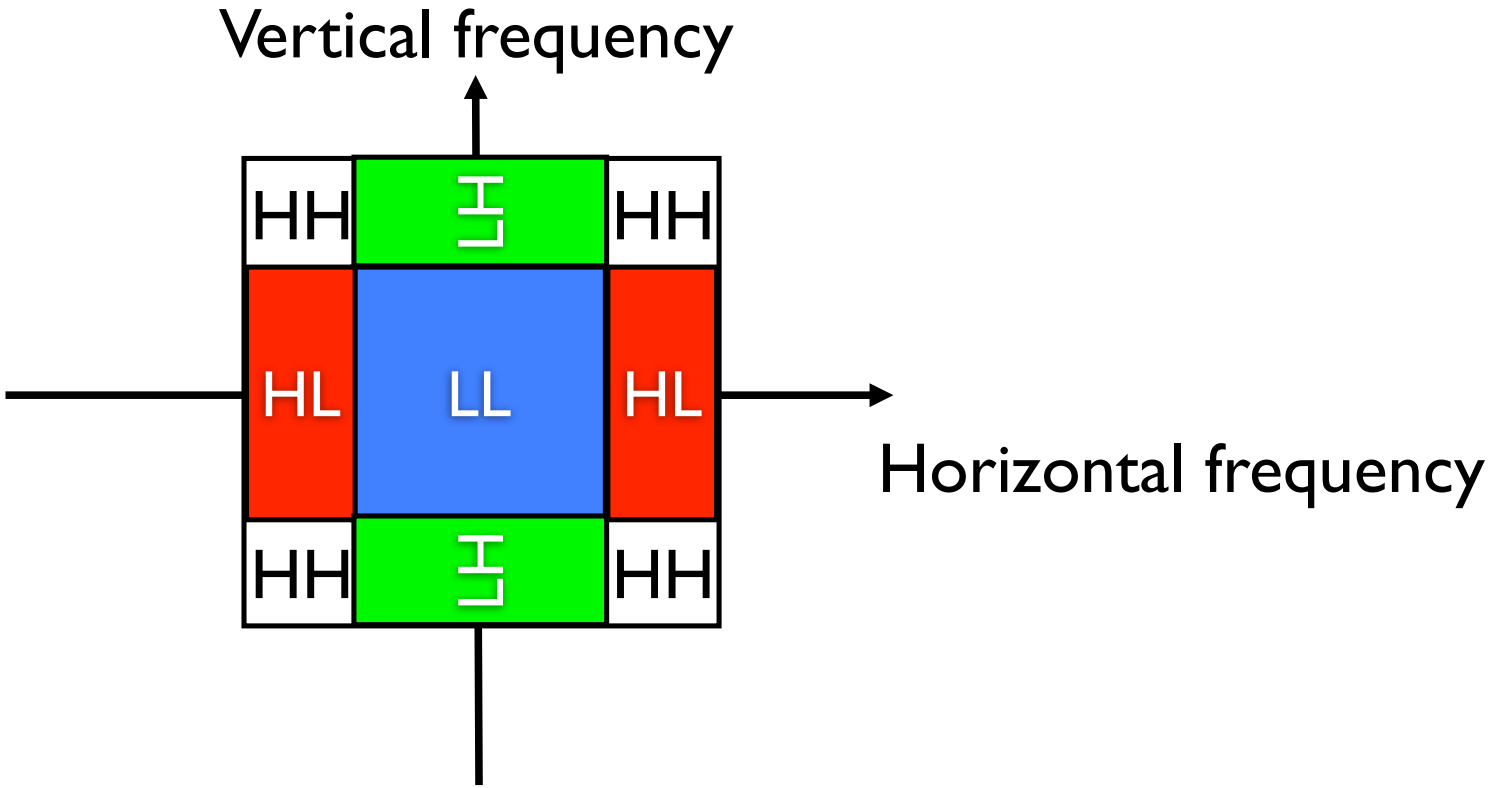




2D separable orthonormal wavelets

- 1D wavelets $\{\psi_{j,k}\}_{j,k} \cup \{\phi_{J,k}\}_k$
- 2D wavelet coefficients
  - LL $\langle x, \phi_{j,k}(p_1)\phi_{j,k}(p_2)\rangle$
  - HL $\langle x, \psi_{j,k}(p_1)\phi_{j,k}(p_2)\rangle$
  - LH $\langle x, \phi_{j,k}(p_1)\psi_{j,k}(p_2)\rangle$
  - HH $\langle x, \psi_{j,k}(p_1)\psi_{j,k}(p_2)\rangle$

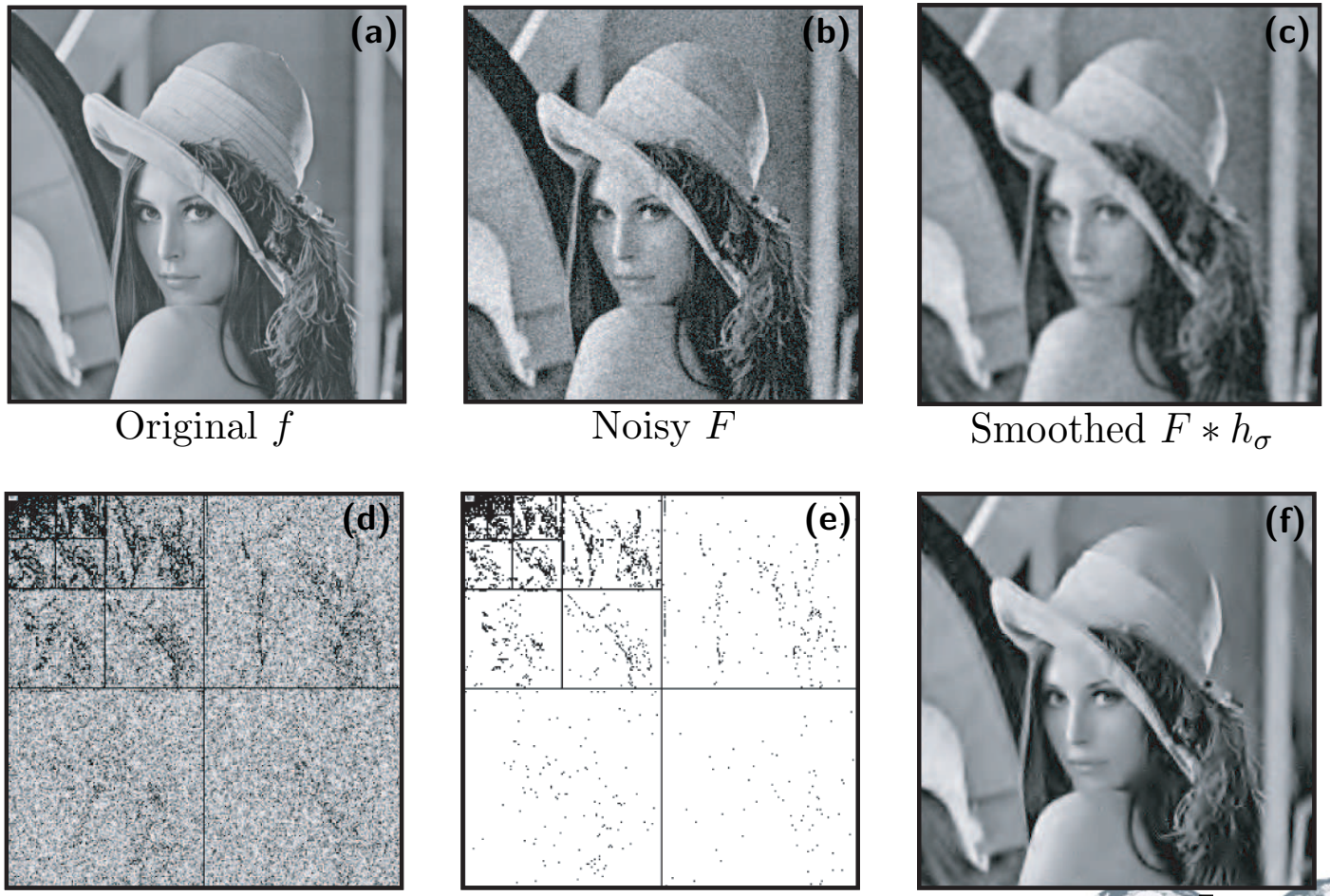
2D wavelets : frequency localization



Wavelet domain processing

Wavelet Domain Denoising

Courtesy: G. Peyré, Ceremade, Université Paris 9 Dauphine

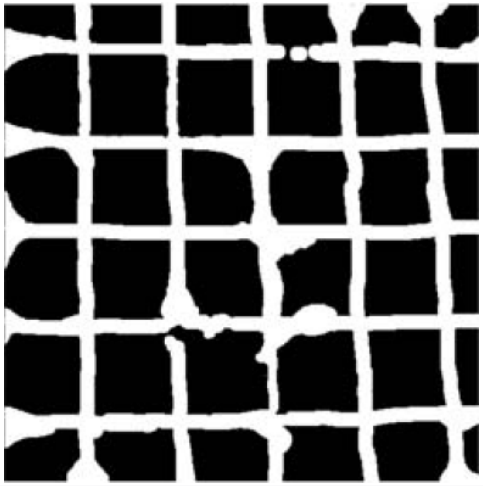


Original f | Noisy F | Smoothed $F * h_\sigma$

Coefficients $\langle f, \psi_m \rangle$ | Thresholded coefficients | Denoised $\bar{f}$

Image inpainting

Courtesy: G. Peyré, Ceremade, Université Paris 9 Dauphine



Inpainting →

Made possible by local texture «extrapolation»

Inpainting model

- Image model

$$x[\vec{p}] = \sum_{j,k} c_{j,k} \psi_{j,k}[\vec{p}]$$

$$x = \mathbf{\Phi} c \in \mathbb{R}^N$$

- Partial observation model

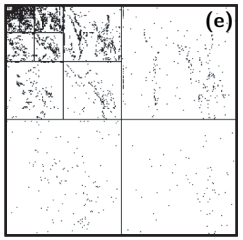
$$y[\vec{p}] = x[\vec{p}], \ \vec{p} \in \text{Observed}$$

$$y[\vec{p}] = 0, \ \vec{p} \in \text{Not observed}$$

$$y = \mathbf{M}x = \mathbf{M}\mathbf{\Phi}c \in \mathbb{R}^m, m < N$$

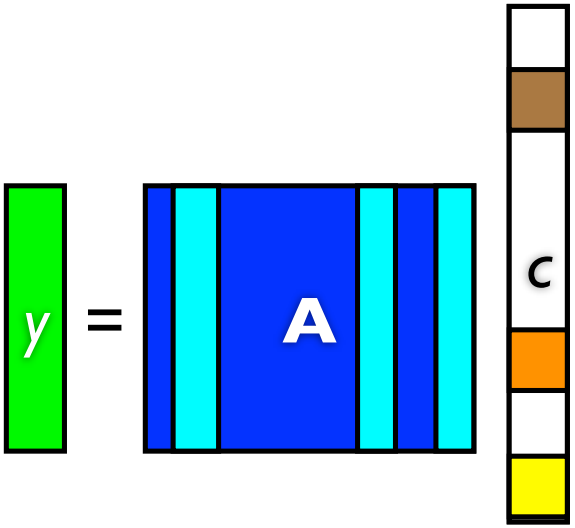
Inpainting problem

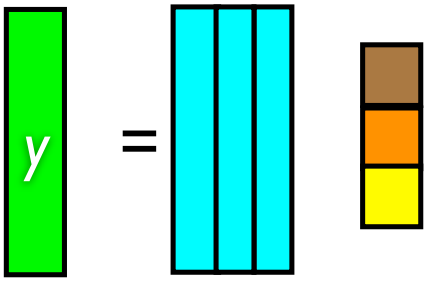
- **Given**: partial observation $y = \mathbf{M}x = \mathbf{M}\mathbf{\Phi}c$
- **Goal**: recover image *x* / representation *c*
- **Issue**: less equations than unknowns
- **Approach**: exploit *sparsity* of *c*



Sparsity

- Under-determined system
- Sparsity:
  - c has few nonzero entries
  - from their location ... recover their value



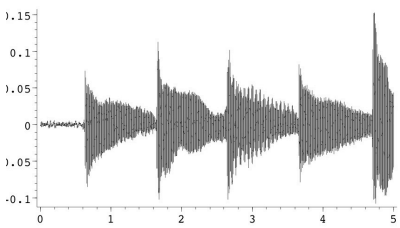


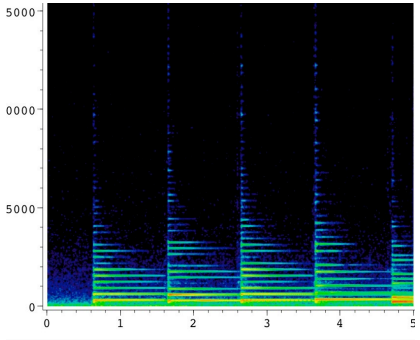
- Finding their location?

Sparse Overcomplete Representations

Examples

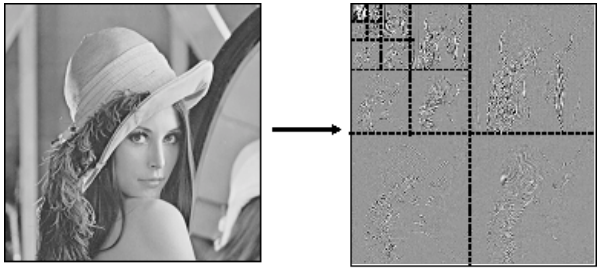
- Locally harmonic : STFT





Black = zero

- Piecewise smooth : wavelets



Grey = zero

Sparse representations

- Fourier / DCT / wavelet orthonormal basis

$$x = \sum_k \langle x, g_k \rangle g_k \in \mathbb{R}^N$$

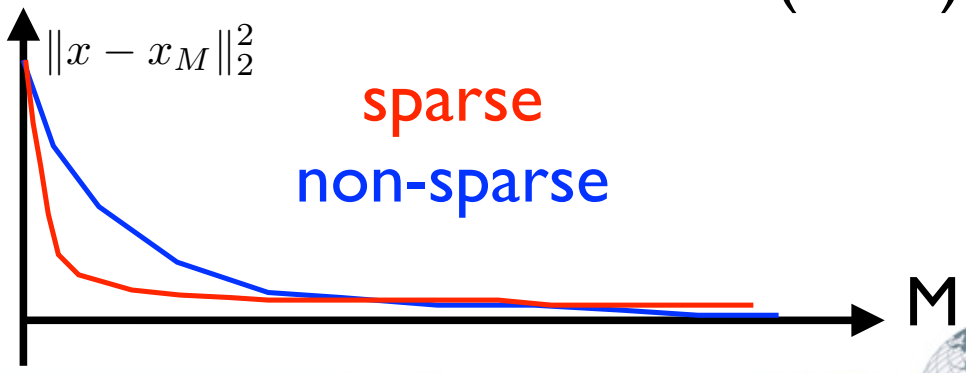
- Notion of *sparsity*
  - few significant coefficients
  - good approximation with few elements
- Related to good compression properties
- Need to adapt basis to type of signals to approximate / compress

Compression & approximation

- Best approximation : keep largest coefficient

$$x = \sum_k \langle x, g_k \rangle g_k \qquad |\langle x, g_{k_1} \rangle| \geq \ldots \geq |\langle x, g_{k_M} \rangle| \geq \ldots \qquad x_M := \sum_{m=1}^{M} \langle x, g_{k_m} \rangle g_{k_m}$$

- Approximation error $\|x - x_M\|_2^2 = \sum_{m \geq M} |\langle x, g_k \rangle|^2$
- Error *vs* number of coefficients (~bits)



Which representation for which signals ?

- Harmonic signals : Fourier basis
- Locally harmonic signals : MDCT (STFT)
- Piecewise smooth signals : wavelets
- Superposition of different “layers” ?
  - Audio = transients + harmonics !
  - Images = contours + textures !

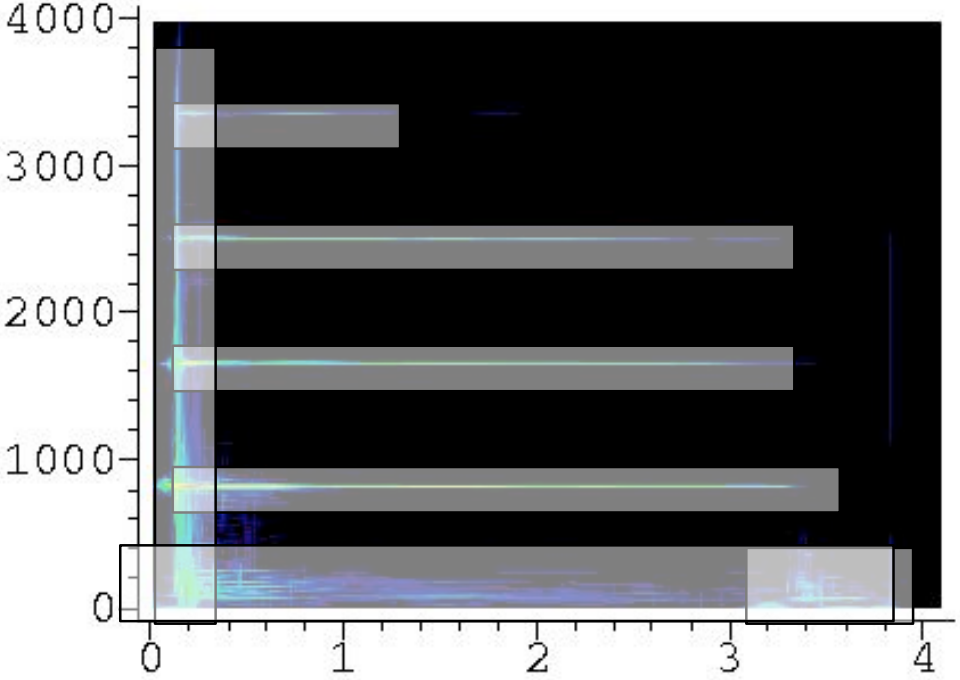
Beyond Heisenberg Uncertainty Principle ?

- A single piano keystroke, 4 sec. @ 8kHz

String = long and tonal

Hammer = transient

Table = low frequency



Adaptive representation with dictionaries

- Time-frequency representation $x(\tau) = \sum_{t,f} X(t,f) w_{t,f}(\tau)$
- Time-scale representations $x(\tau) = \sum_{j,k} \langle x, \psi_{j,k} \rangle \psi_{t,s}(\tau)$
- General model :
  - redundant *dictionary* of *atoms* $g_k(t)$
  - linear synthesis $x(t) \approx \sum_k a_k g_k(t)$
- Objective : compute representation $\{a_k\}$

Redundant dictionaries

- There are more basis vectors than the signal dimension $\{g_k[n]\}_{k=1}^{K}$ $\quad x = (x[n])_{n=1}^{N}, N < K$
- There are infinitely many possible representations

$$x = \sum_k c_k g_k = \sum_k d_k g_k \Longleftrightarrow \sum_k (c_k - d_k) g_k = 0$$

- Preferably choose *sparse* representations

Greedy algorithms

- Idea : the most significant components are associated to large inner products
- *Matching Pursuit* = iterative greedy algorithm
  - pick the best atom $k_i := \arg\max_k |\langle x^{(i)}, g_k \rangle|$
  - remove it from the residual $x^{(i+1)} := x^{(i)} - \langle x^{(i)}, g_{k_i} \rangle g_{k_i}$

  after M steps $x = \sum_{i=0}^{M-1} \langle x^{(i)}, g_{k_i} \rangle g_{k_i} + x^{(M)}$
- Orthogonal Matching Pursuit : $x^{(i+1)} := x^{(i)} - \sum_{l=0}^{i-1} \alpha_l^{(i)} g_{k_i}$

  (Mean Square Error projection)

Summary: key concepts

- Convolution: definition, main properties
- Fourier transform: definition, main properties
- Fast convolution: overlap-add algorithm
- Short-time Fourier transform: definition, role of the window, reconstruction
- Wavelet transform: principle, 1D & 2D
- Notion of sparsity: role in data compression