Aisenstadt Chair Course
CRM September 2009

Part III
Super-Resolution with Sparsity

Stéphane Mallat

Centre de Mathématiques Appliquées
Ecole Polytechnique

Super-Resolution with Sparsity

- ***Dream:*** recover high-resolution data from low-resolution noisy measurements:
  - Medical imaging
  - Satellite imaging
  - Seismic exploration
  - High Definition Television or Camera Phones
- Can we improve the signal resolution ?
- Sparsity as a tool to incorporate prior information.

Inverse Problems

- Measure a noisy and low resolution signal:

$$Y = Uf + W$$

with $f \in \mathbf{C}^N$ and $\dim(\mathbf{ImU}) = Q < N$.

- Inverse problems: compute an estimation

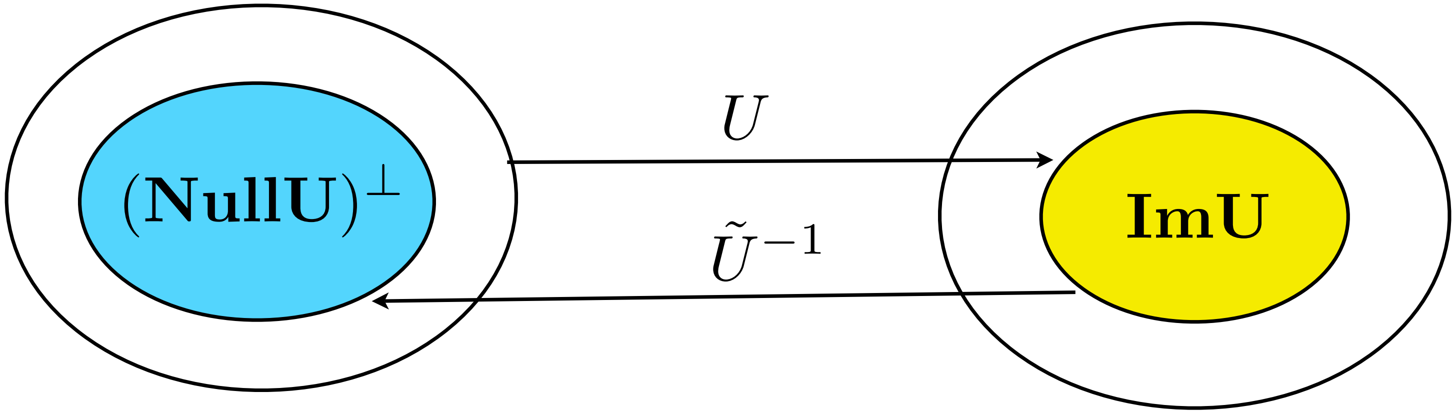
$$\tilde{F} = DY$$

and minimize the risk: $r(D, f) = E\{\|\tilde{F} - f\|^2\}$

- Super-resolution estimation: $\tilde{F}$ is computed in a space of dimension □ Is it possible, how ?

Regularized Inversion

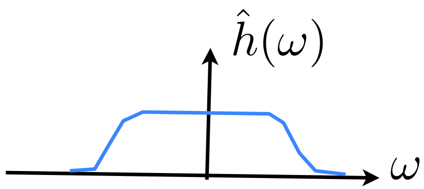
To estimate f from $Y = Uf + W$ invert U !



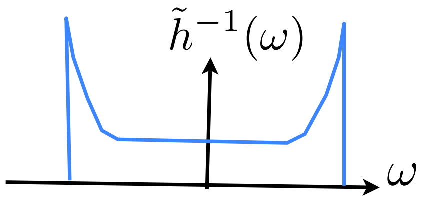
Pseudo inverse: $\tilde{U}^{-1} U f = f$ if $f \in (\mathbf{NullU})^\perp$

$\tilde{U}^{-1} g = 0$ if $g \in (\mathbf{ImU})^\perp$

Deconvolution: $Uf = f * h$ with



$\tilde{U}^{-1} f = f * \tilde{h}^{-1}$ with



Regularization and Denoising

$$\tilde{U}^{-1} Y = \tilde{U}^{-1} U f + \tilde{U}^{-1} W$$

Problems: $\tilde{U}^{-1} U f \in (\mathbf{Null U})^{\perp}$ no super-resolution

$\|\tilde{U}^{-1} W\|$ is huge if $\tilde{U}^{-1}$ is not bounded.

Regularized inversion includes a noise reduction with a projection in a space $\mathbf{V}$:

$$\tilde{F} = R(\tilde{U}^{-1} Y) \in \mathbf{V}$$

Optimizing *R* requires prior information.

No super-resolution : $\dim(\mathbf{V}) \leq Q$.

Singular Value Decompositions

- Basis of singular vectors $\{e_k\}_{1\leq k\leq N}$ diagonalizes U^*U:

$$U^*U e_k = \lambda_k^2\, e_k$$

- Diagonal denoising over the singular basis:

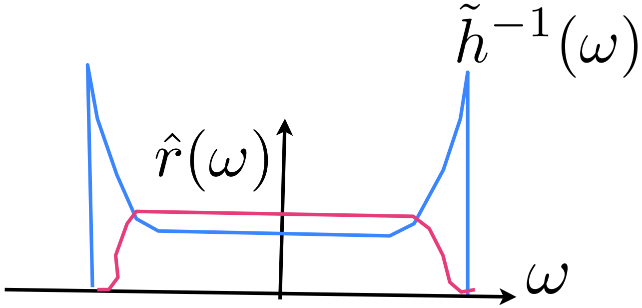
$$\tilde{F} = R(\tilde{U}^{-1}Y) = \sum_{k=0}^{N-1} r_k\, \langle \tilde{U}^{-1}Y, e_k\rangle\, e_k\ .$$

$$\text{Since}\quad \langle \tilde{U}^{-1}Y, e_k\rangle = \lambda_k^{-2}\, \langle Y, U e_k\rangle$$

$$r_k = \frac{1}{1+\sigma^2\,\lambda_k^{-2}} \quad \text{yields} \quad \tilde{F} = \sum_{k=0}^{N-1} \frac{\langle Y, U e_k\rangle}{\lambda_k^2+\sigma^2}\, e_k\ .$$

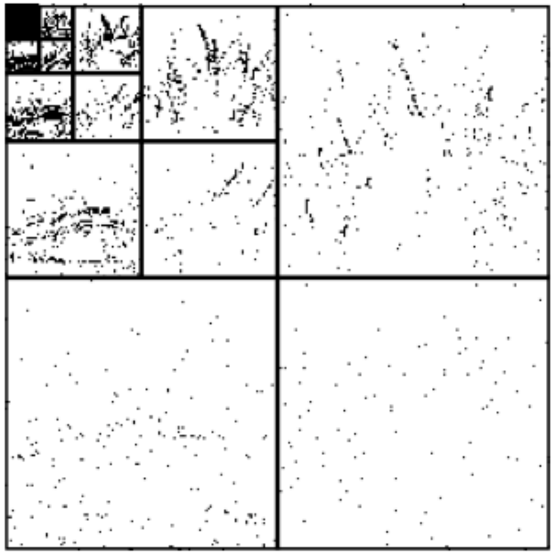
Linear filtering of deconvolution:

$$R(\tilde{U}^{-1}Y) = \tilde{U}^{-1}Y * r$$



Denoising by Thresholding

Non linear projector adapted to the signal:


Threshold $T = 3\,\sigma$ where σ^2 is the noise variance.

Thresholding for Inverse Problems

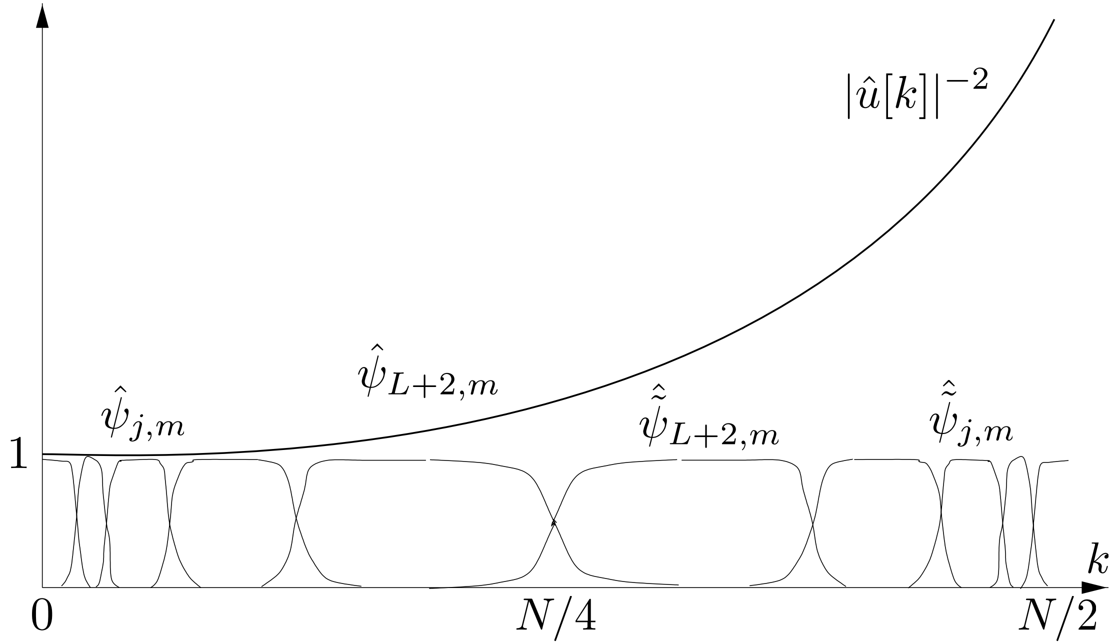
- Remove noise from $\tilde{U}^{-1}Y = \tilde{U}^{-1}Uf + \tilde{U}^{-1}W$ with a thresholding estimator.

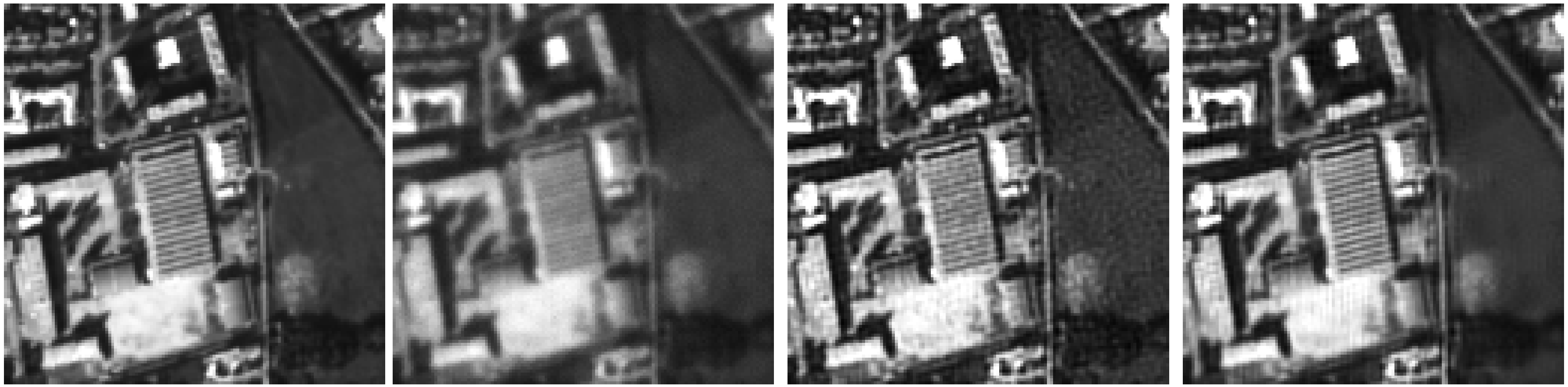
- Optimal in a basis $\{\phi_p\}_{p\in\Gamma}$ providing a sparse representation of f and which decorrelates the noise coefficients $\langle \tilde{U}^{-1}W, \phi_p \rangle$.

- The dictionary vectors ϕ_p must be almost eigenvectors of U^*U , they must have a narrow spectrum:

$$\phi_p = \sum_{k\in S_p} \langle \phi_p, e_k \rangle \, e_k \quad \text{with} \quad \lambda_k^2 \sim \tilde{\lambda}_p^2 \text{ for } k \in S_p$$

Satellite Image Deconvolution



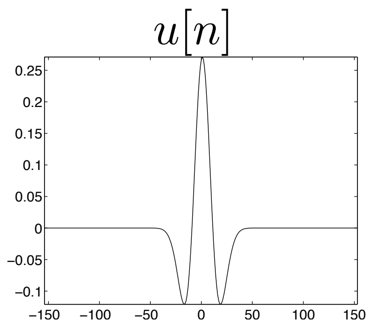


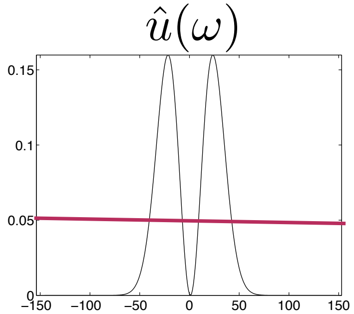
Original Smoothed Linear Wiener Thresholding

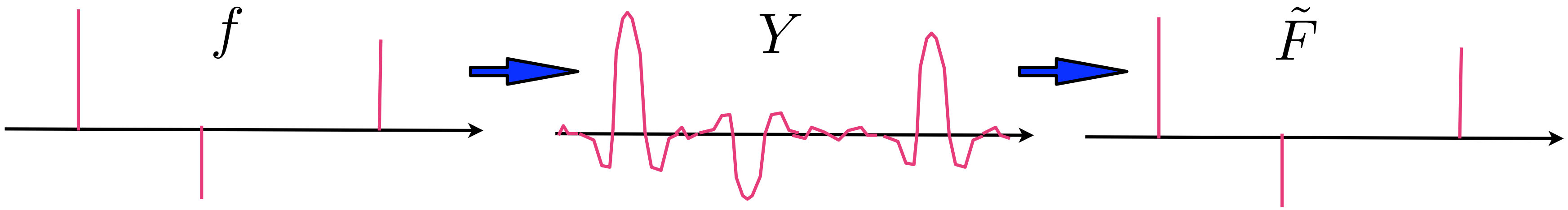
Sparse Spike Deconvolution

Seismic data: $Y = f * u + W$ with $f[n] = \sum_{p \in \Lambda} a[p]\, \delta[p-n]$

$$Y[q] = \sum_{p \in \Lambda} a[p]\, u[q-n] + W[q]$$







Super-resolution inversion by detection of the sparse support

Sparse Super-Resolution

- ***Prior information:*** f has a sparse approximation in a normalized dictionary $\mathcal{D} = \{\phi_p\}_{p\in\Gamma}$ of at least N vectors

$$f = \sum_{p\in\Lambda} a[p]\,\phi_p + \epsilon_\Lambda$$

with a small error $\|\epsilon_\Lambda\|$.

It results that

$$Y = Uf + W = \sum_{p\in\Lambda} a[p]\,U\phi_p + \ (U\epsilon_\Lambda + W)$$

has a sparse approximation in the ***redundant dictionary***

$$\mathcal{D}_U = \{U\phi_p\}_{p\in\Gamma}$$

in the space $\mathbf{ImU}$ of dimension $Q \leq N$

Sparse Super-Resolution

- A sparse approximation of Y is computed in $\mathcal{D}_U = \{U\phi_p\}_{p\in\Gamma}$

$$Y_{\tilde{\Lambda}} = \sum_{p\in\tilde{\Lambda}} \tilde{a}[p]\, U\phi_p$$

with a pursuit algorithm. A basis pursuit minimizes the Lagrangian:

$$\|Y - \sum_{p\in\Gamma} \tilde{a}[p]\, U\phi_p\|^2 + \lambda \sum_{p\in\Gamma} |\tilde{a}[p]|$$

and $\tilde{\Lambda}$ is the support of $\tilde{a}$.

- It yields a signal estimator $\tilde{F} = \sum_{p\in\tilde{\Lambda}} \tilde{a}[p]\, \phi_p$

using prior information which recovers ϕ_p from each $U\phi_p$.

Error and Exact Recovery

- From the sparse decomposition of $Y = f + W$

$$Y_{\tilde{\Lambda}} = \sum_{p \in \tilde{\Lambda}} \tilde{a}[p] \, U \phi_p$$

since $$f = \sum_{p \in \Lambda} a[p] \, \phi_p + \epsilon_\Lambda$$

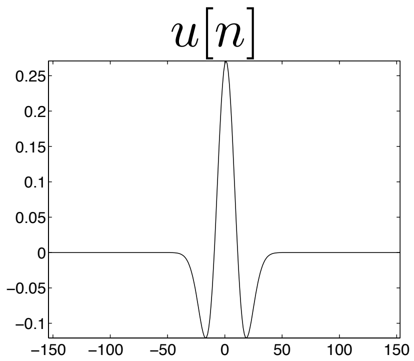
$$\|f - \tilde{F}\| \leq \left\| \sum_{p \in \tilde{\Lambda}} \tilde{a}[p] \phi_p - \sum_{p \in \Lambda} a[p] \, \phi_p \right\| + \|\epsilon_\Lambda\| \ .$$

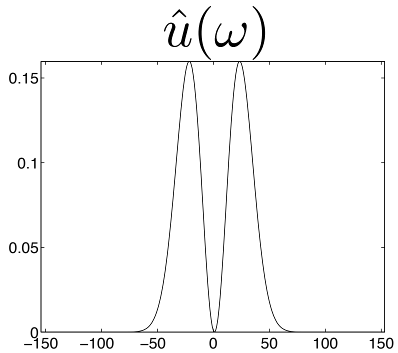
- Small error if $\tilde{\Lambda}$ includes Λ and if $\{U\phi_p\}_{p \in \tilde{\Lambda}}$ is a Riesz basis.
- Exact recovery in the redundant dictionary $\mathcal{D}_U = \{U\phi_p\}_{p \in \Gamma}$
- Super-resolution: if Λ is not restricted to a space of dimension Q.

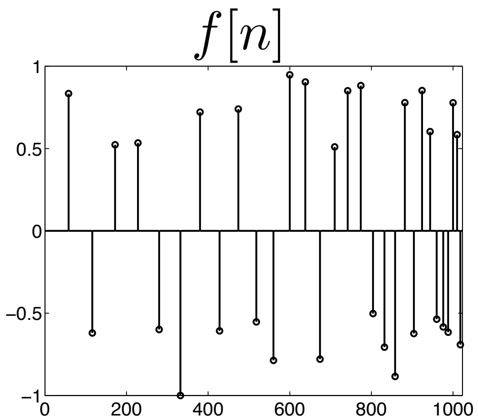
Sparse Spike Deconvolution

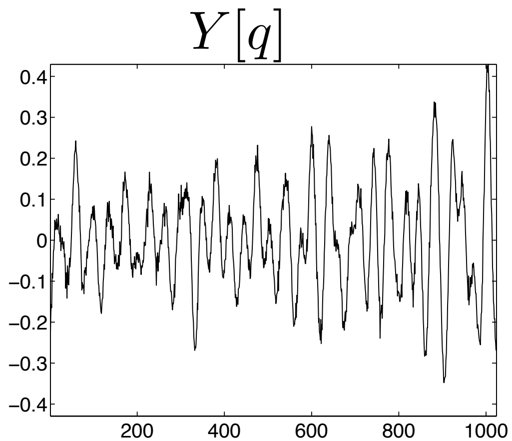
Seismic data: $Y = f * u + W$ with $f[n] = \sum_{p \in \Lambda} a[p]\, \delta[p-n]$

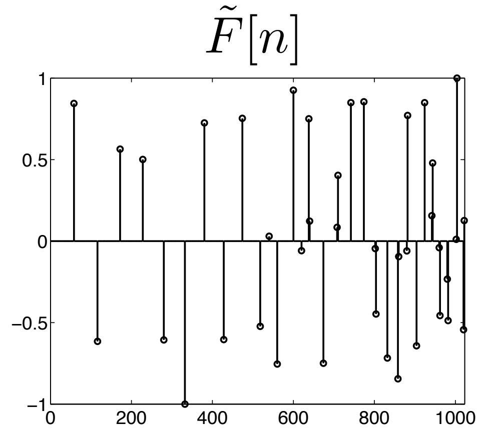
$$\phi_p[n] = \delta[p-n] \quad , \quad U\phi_p[q] = u[q-n] \; , \; \tilde{F}[n] = \sum_{p \in \tilde{\Lambda}} \tilde{a}[p]\, \delta[n-n]$$





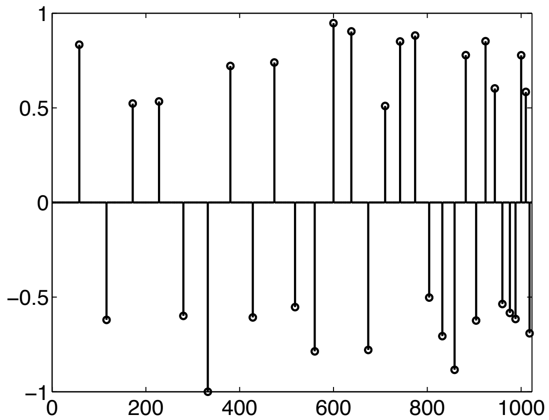




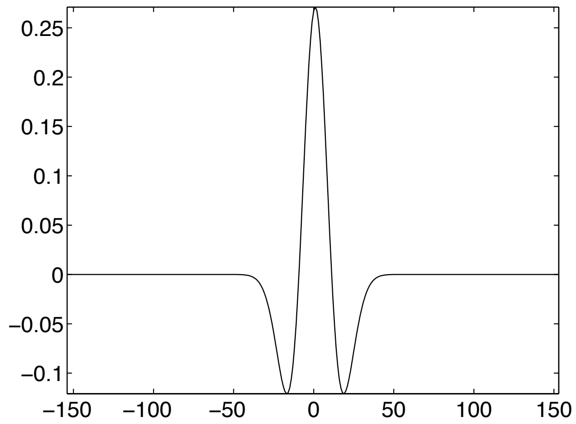


Comparison of Pursuits

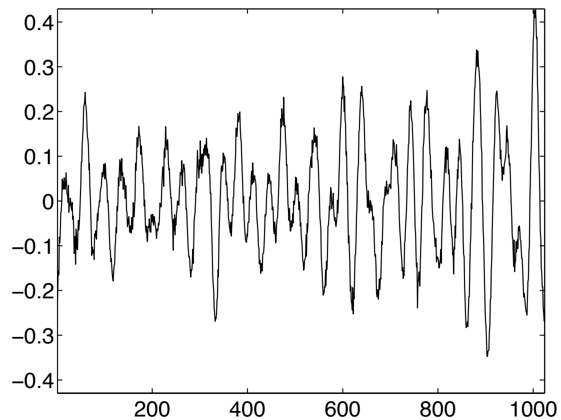
Original



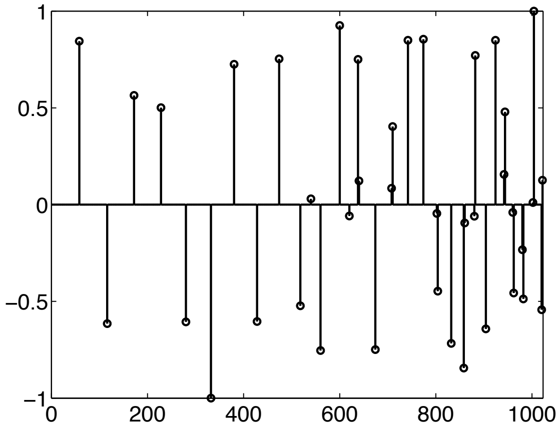
Seismic wavelet



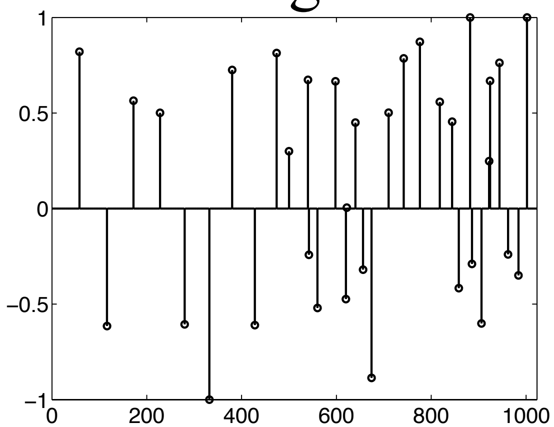
Seismic trace



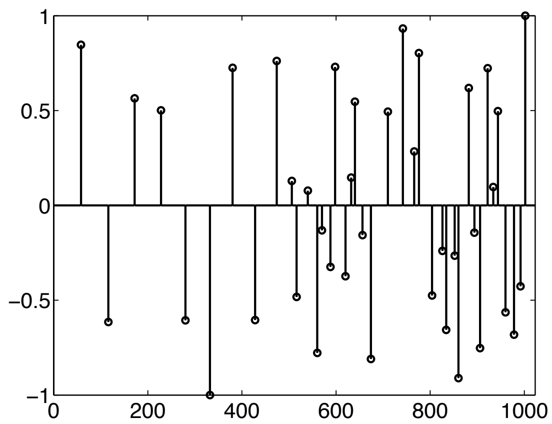
Basis Pursuit



Matching Pursuit



Orth. Match. Pursuit



Conditions for Super-resolution

- The signal approximation support Λ should small.
- Stability: $\{U\phi_p\}_{p\in\Lambda}$ must be a Riesz basis

  $\|U\phi_p\|$ should not be too smal.
- Hence the ϕ_p must have a "spread spectrum" relatively to U^*U.
- Support recovery: the dictionary $\mathcal{D}_U = \{U\phi_p\}_{p\in\Gamma}$ must be as incoherent as possible.
- Exact recovery criteria: $ERC(\Lambda) < 1.$

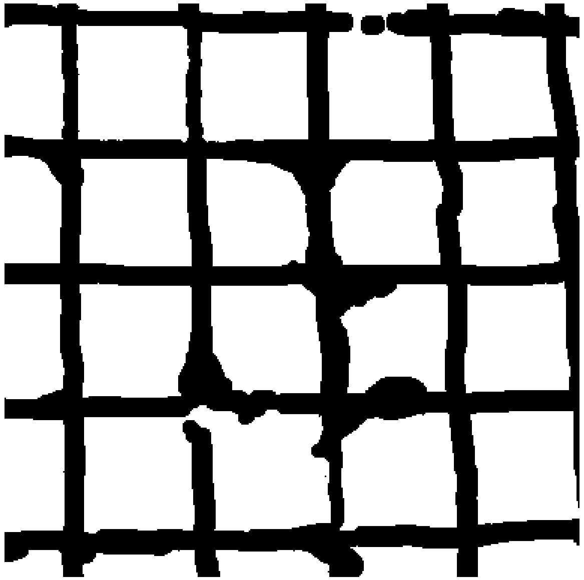
Image Inpainting

$$U f[q] = f[q] \quad \text{for} \quad q \in \Omega \quad \text{with} \quad |\Omega| = Q < N$$

Super-resolution in a wavelet dictionary $\mathcal{D}_U = \{U \phi_p\}_{p \in \Gamma}$

Original

Support of Ω



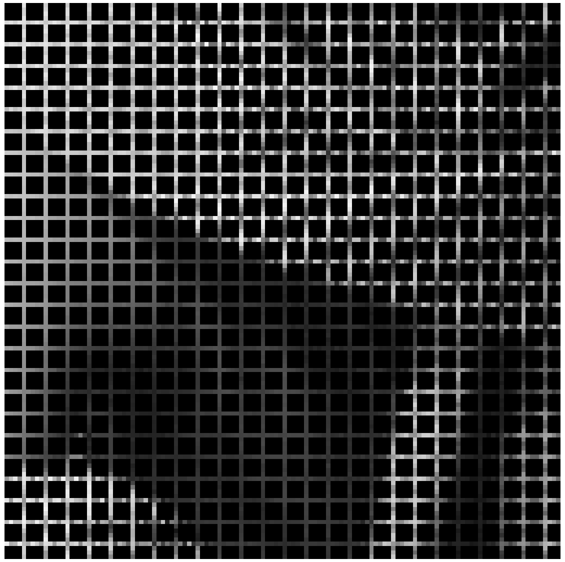
Super-resolution

Image Inpainting

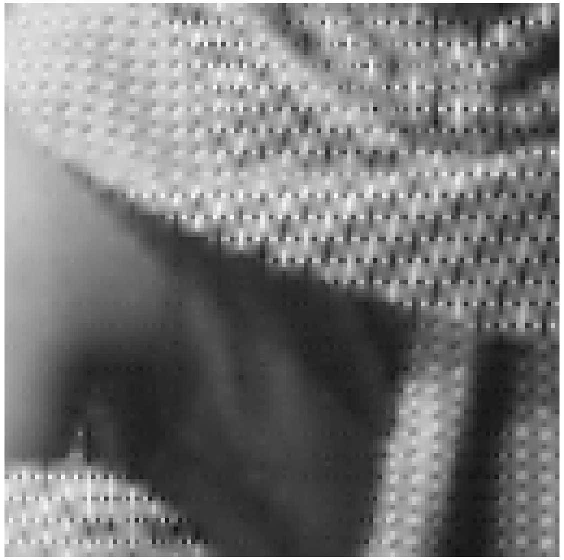
$$Uf[q] = f[q] \quad \text{for} \quad q \in \Omega \quad \text{with} \quad |\Omega| = Q < N$$

Wavelet and local cosinedictionary $\mathcal{D}_U = \{U\phi_p\}_{p\in\Gamma}$

$Y = Uf + W$



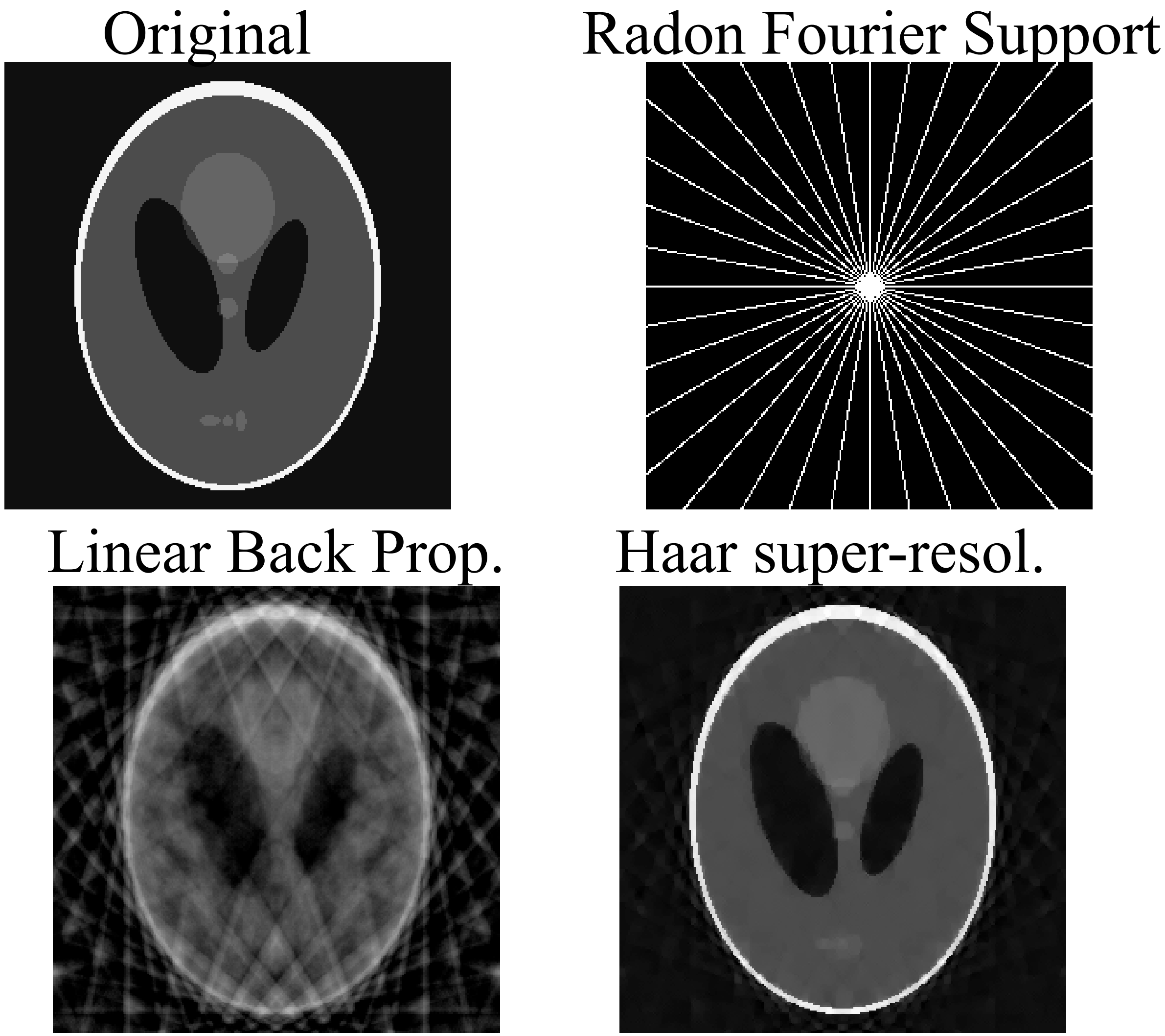
Linear estimation



Super-resolution

Tomography

U is a Radon transform which integrates along straight lines.

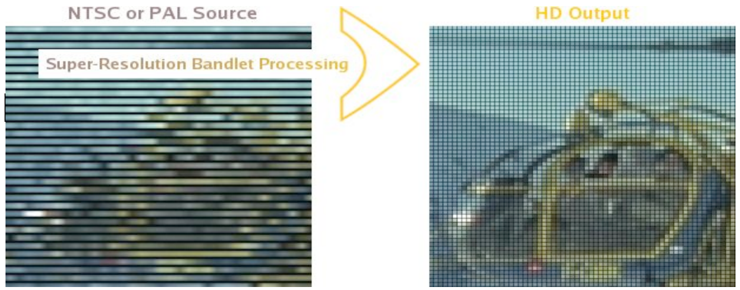


Super-Resolution Zooming

- Need to increase numerically acquired image resolution:
  - Conversion to HDTV of SDTV, Internet and Mobile videos...

Size increase:

60 images of 720 x 576 pixels = 320



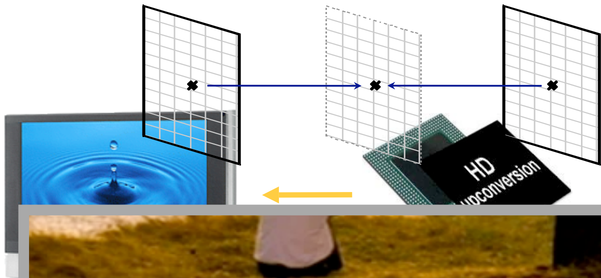
- Spatial deinterlacing and up-scaling
  - **up to 8 times more pixels**

PAL/NTSC

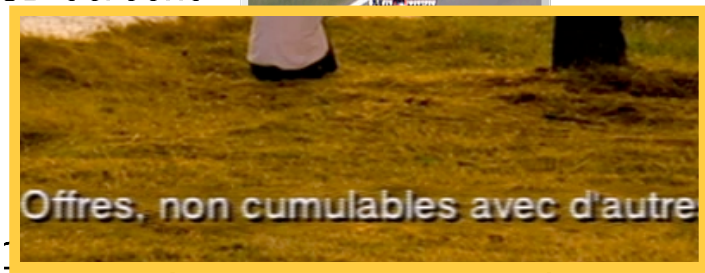
> Vidéo processor in the TV :

x 20

- Frame rate conversion
  - **twice more images for LCD screens**



HD LCD screens



7.5 G

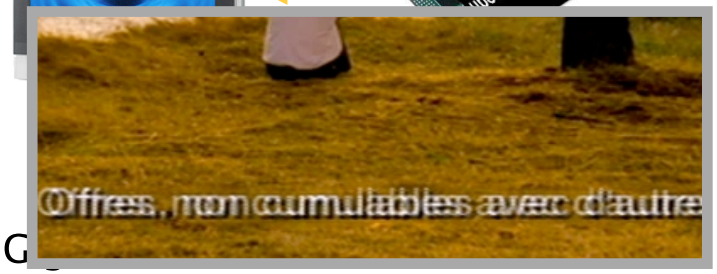


Image and Video Zooming

- Image subsampling : $Uf = f[n/s]$ is a linear projector.
- Linear inversion without noise: linear interpolation

$$\tilde{f}[p] = \sum_n Uf[n]\,\theta(p - ns)$$

- Prior information: geometric regularity.
- Super-resolution by interpolations in the directions of regularity
- Sparse super-resolution marginally improves linear interpolations.

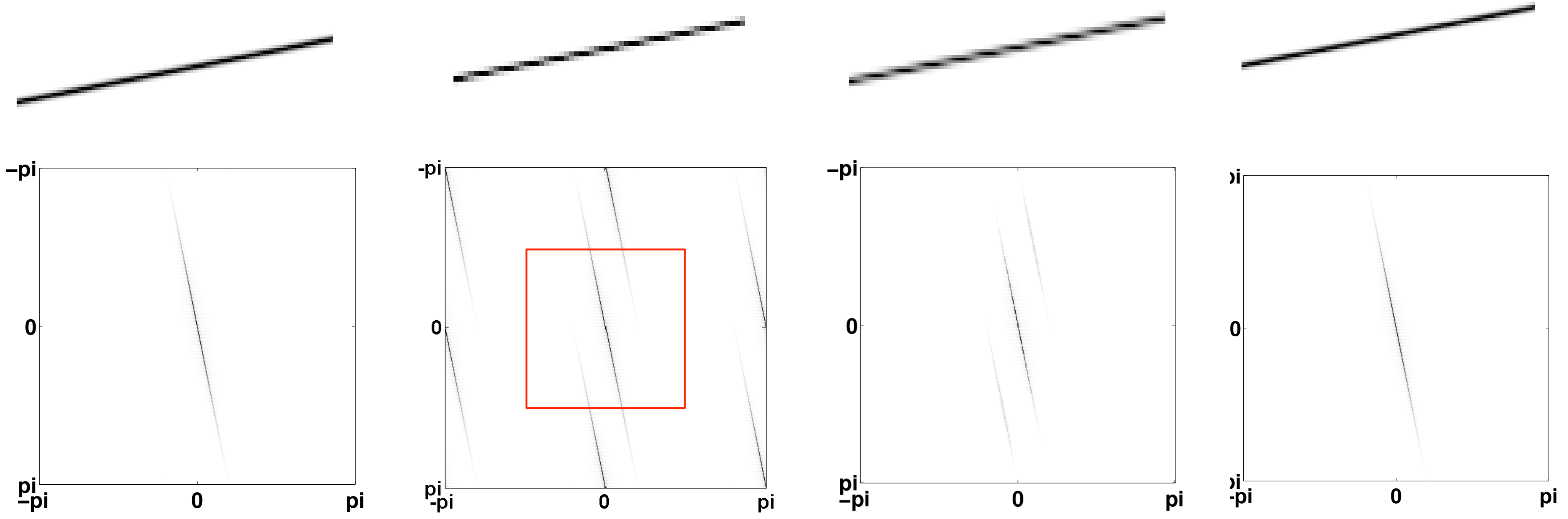
High Resolution Image

Low Resolution Image

Contourlet SuperResolution 28.59 db

Cubic spline Interpolation 28.47 db

Aliased Interpolation



Original Subsampled Linear Interp. Direct. Interp.

Super-resolution is not possible for horizontal and vertical edges.

Adaptive Directional Interpolations

- Linear Tikhonov estimation: $\tilde{F}_\theta = I_\theta Y$

$$\text{minimizes } \|R_\theta I_\theta Y\| \text{ subject to } U I_\theta Y = Y$$

where R_θ is a linear directional regularity operator.

- Adaptive directional interpolation adapt locally θ by testing locally the directional regularity with gradient operators.

- General class of mixing linear operators in a frame $\{\phi_p\}_{p\in\Gamma}$

$$\tilde{F} = \sum_{\theta\in\Theta} I_\theta \left(\sum_{p\in\Gamma} a(\theta,p)\, \langle Y, \phi_p\rangle\, \phi_p \right)$$

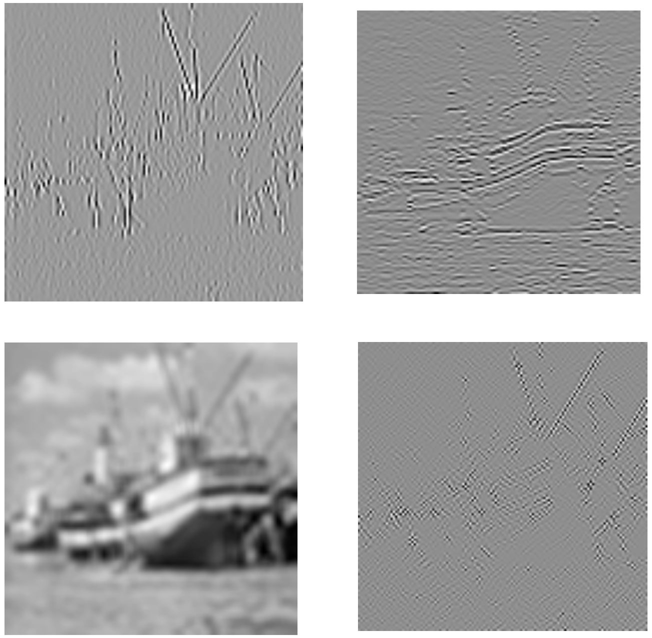
- Problem: how to optimize the $a(\theta, p)$?

Wavelet Block Interpolation

Wavelet transform on 1 scale, $j = 1$

$$\langle f, \psi_{j,n}^k \rangle = \int f(x)\, 2^{-j}\, \psi^k(2^{-j}(x-n))\, dx$$

$k = 1$ $k = 2$



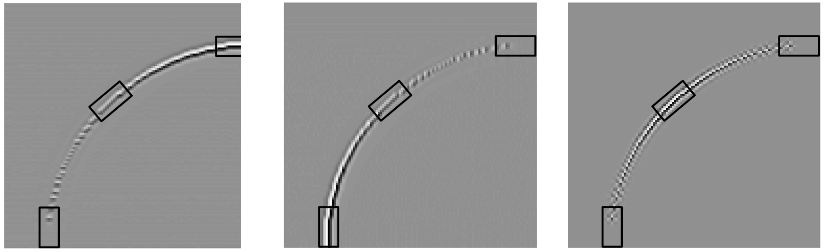
$2^j = 2$

$k = 3$

Low frequencies are linearly interpolated (no aliasing).
Adaptive directional interpolation of fine scale wavelets.

Wavelet Block Interpolation

- Dictionary of blocks $\{B_{\theta,q}\}_{\theta,q}$



- To a wavelet block decomposition

$$Y = \sum_{\theta} \sum_{q} \epsilon(\theta, q)\, P_{B_{\theta,q}} Y + Y_r$$

with $$P_{B_{\theta,q}} Y = \sum_{(n,k) \in B_{\theta,q}} \langle Y, \psi^k_{1,n} \rangle\, \psi^k_{1,n}$$

we associate an interpolation estimation

$$\tilde{F} = \sum_{\theta} I_\theta \left(\sum_{q} \epsilon(\theta, q)\, Y_{q,\theta} \right) + I_r(Y_r)$$

- How to optimize the $\epsilon(\theta, q)$?

Adaptive Tikhonov Estimation

- To compute

$$Y = \sum_{\theta} \sum_{q} \epsilon(\theta, q)\, P_{B_{\theta,q}} Y + Y_r$$

where $\epsilon(\theta, q)$ is sparse and $\epsilon(\theta, q) \approx 1$ if $\|R_\theta I_\theta P_{B_{\theta,q}} Y\|$ is small: Lagrangian minimization

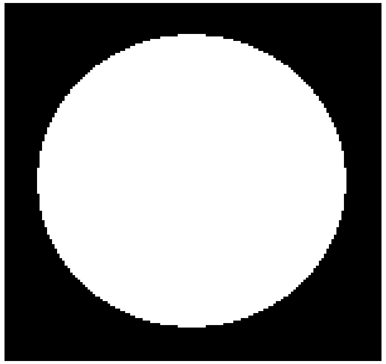
$$\mathcal{L} = \|Y - \sum_{\theta,q} \epsilon(\theta, q)\, P_{B_{\theta,q}} Y\|^2 + \lambda \sum_{\theta,q} |\epsilon(\theta, q)|\, \|R_\theta I_\theta P_{B_{\theta,q}} Y\|^2$$

- Standard $\mathbf{l}^1$ minimization. Can be solved with a greedy pursuit.
- If there is only one $\epsilon(\theta, q) \neq 0$ then $\mathcal{L}$ is minimized by

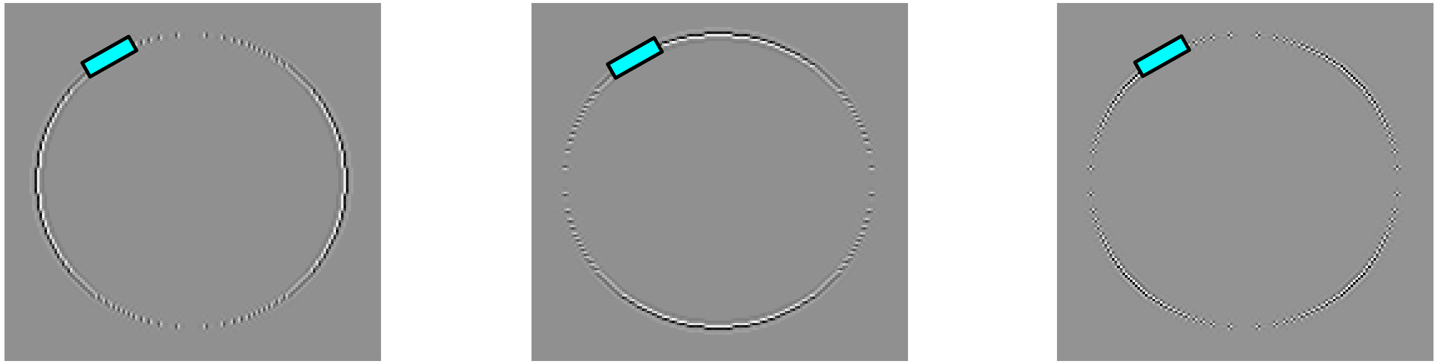
$$\epsilon(\theta, q) = \max\left(1 - \lambda \frac{\|R_\theta I_\theta P_{B_{\theta,q}} Y\|^2}{\|P_{B_{\theta,q}} Y\|^2}, 0\right) \quad \text{and}$$

$$\mathcal{L} = \|Y\|^2 - e(\theta, q) \quad \text{with} \quad e(\theta, q) = \frac{\|P_{B_{\theta,q}} Y\|^2\, \epsilon(\theta, q)^2}{2} \,.$$

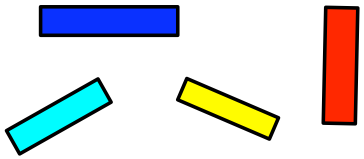
Wavelet Block Spaces



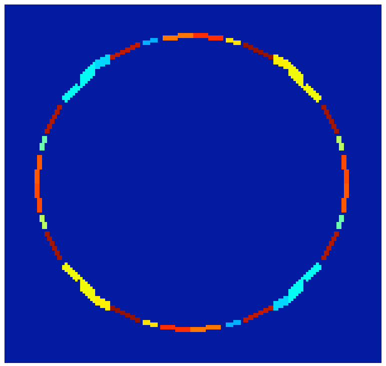
Wavelet transform



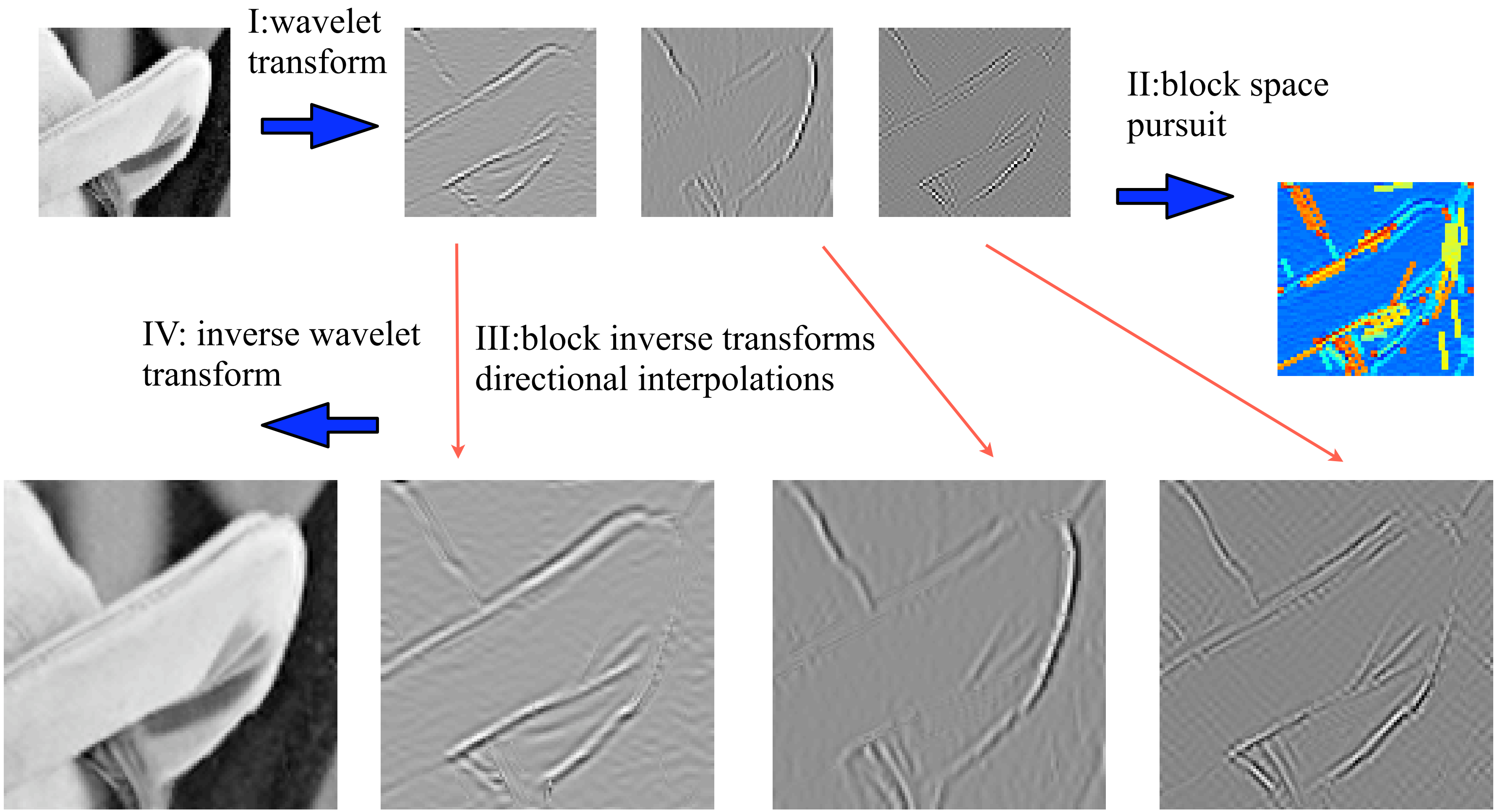
Blocks of oriented bars



Block projection pursuit

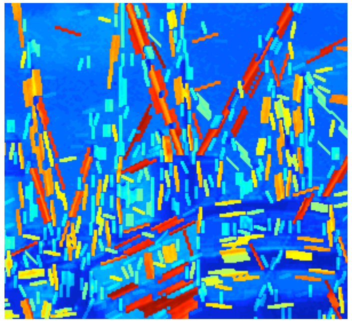


Super-Resolution Block Interpolation



Comparison with Cubic Splines

Block pursuit
on wavelet coefficients

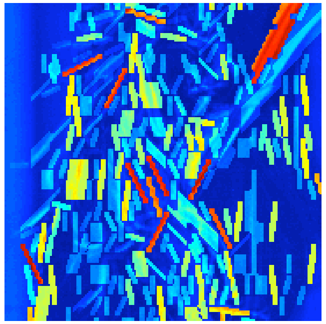


Block Interpolations
over wavelet coefficients

Cubic spline interpolations

Comparison with Cubic Splines

Block pursuit
on wavelet coefficients

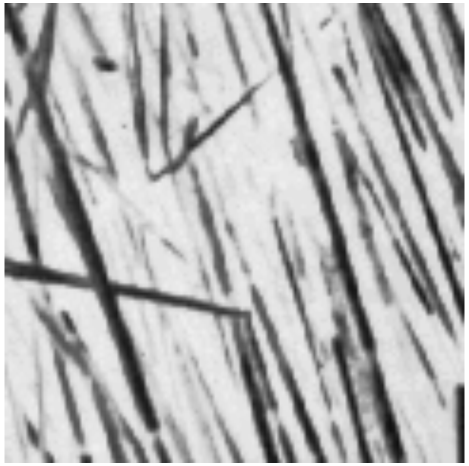


Block Interpolations SNR = 29.24 db
over wavelet coefficients

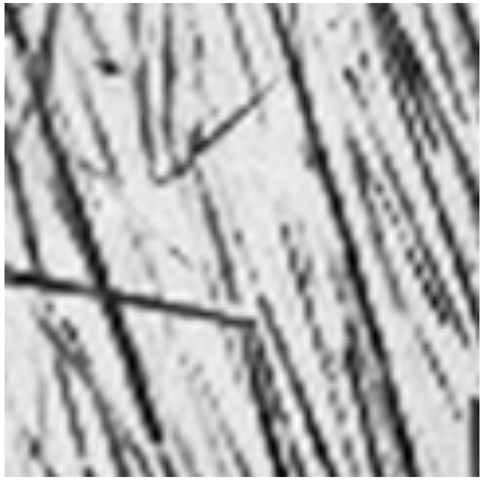
SNR = 28.58 db
Cubic spline interpolations

Examples of Zooming

Original Image

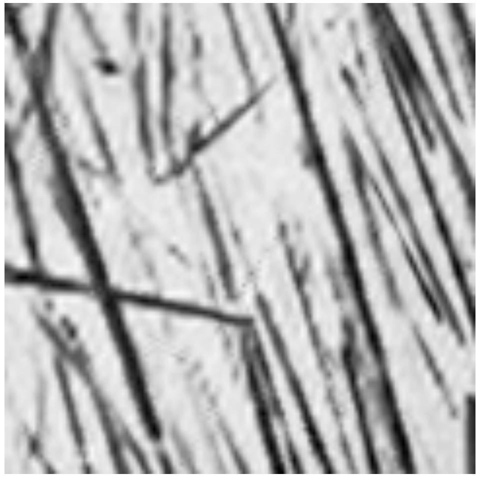


Cubic Spline Interpolation



SNR = 22.35 db

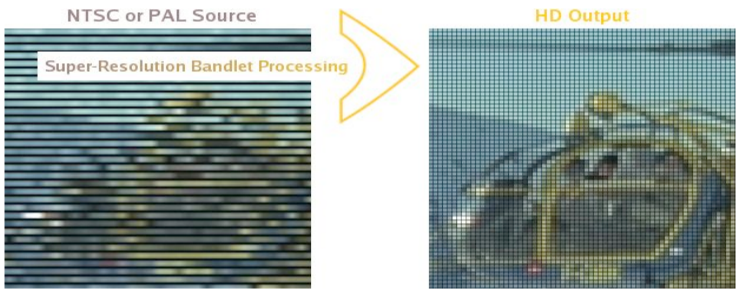
Bandlet Super-Resolution



SNR = 24.14 db

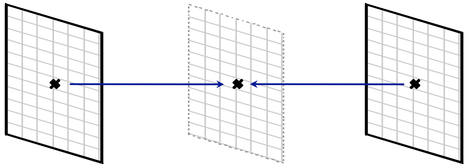
Super-Resolution Zooming

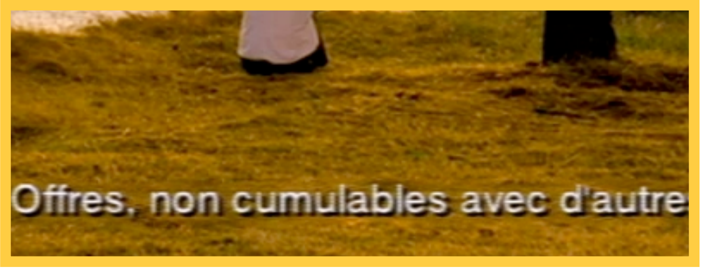
- Need to increase numerically acquired image resolution:
  - Conversion to HDTV of SDTV, Internet and Mobile videos...

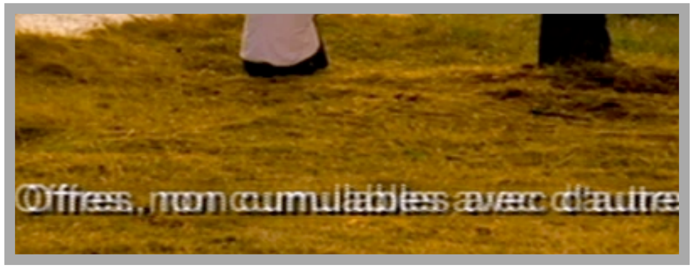


- Spatial deinterlacing and up-scaling
  - up to 8 times more pixels

- Frame rate conversion
  - twice more images for LCD screens







3rd. Concluion

- Super-resolution is possible for signals that are sparse in a dictionary $\mathcal{D} = \{\phi_p\}_{p\in\Gamma}$ which has a spread spectrum and which is transformed in an incoherent dictionary $\mathcal{D}_U = \{U\phi_p\}_{p\in\Gamma}$

- Super-resolution is typically not possible for any class of signals

- Need to incoporate as much prior informaiton as possible: use of structured sparse representations.

- What if it was possible to choose the operator U ?

  compressed sensing...