Fourier, Wavelets and Beyond: The Search for Good Bases for Images

Martin Vetterli
EPFL & UC Berkeley

Colloque Jacques Morgenstern
7.06.07 INRIA Sophia-Antipolis

Audiovisual Communications Laboratory

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Acknowledgements

Collaborations:

- T.Blu, EPFL
- M.Do, UIUC
- P.L.Dragotti, Imperial College
- P.Marziliano, NTU Singapore
- C.Weidmann, TRC Vienna

Discussions and Interactions:

- A.Cohen, Paris VI
- I. Daubechies, Princeton
- R.DeVore, Carolina
- D. Donoho, Stanford
- M.Gastpar, Berkeley
- V.Goyal, MIT
- J. Kovacevic, CMU
- S. Mallat, Polytech. & NYU
- M.Unser, EPFL

Sponsors: NSF Switzerland

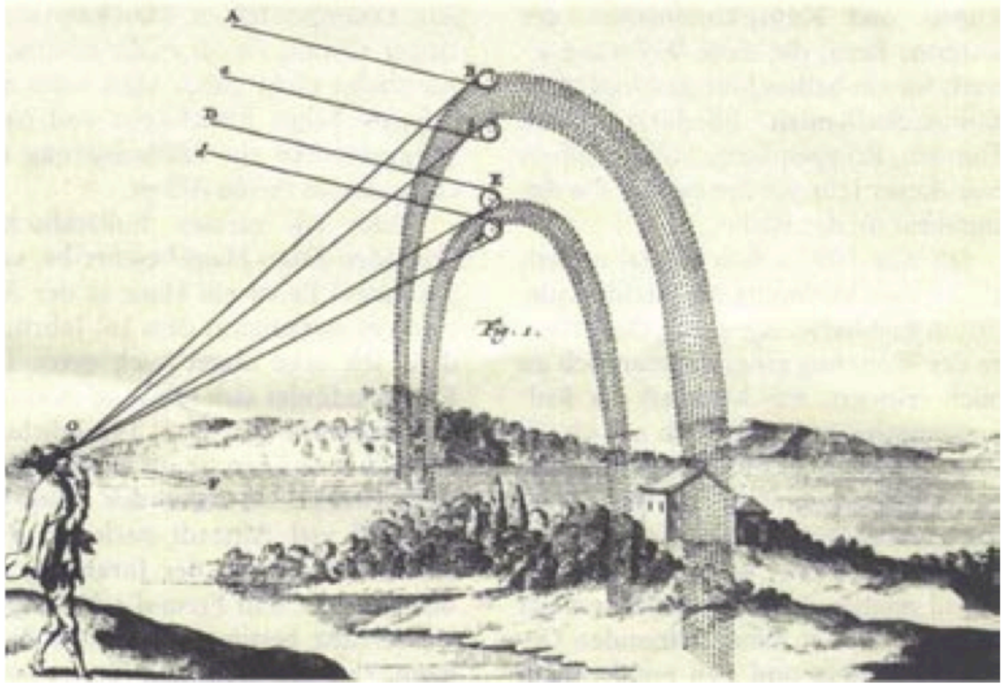
Outline

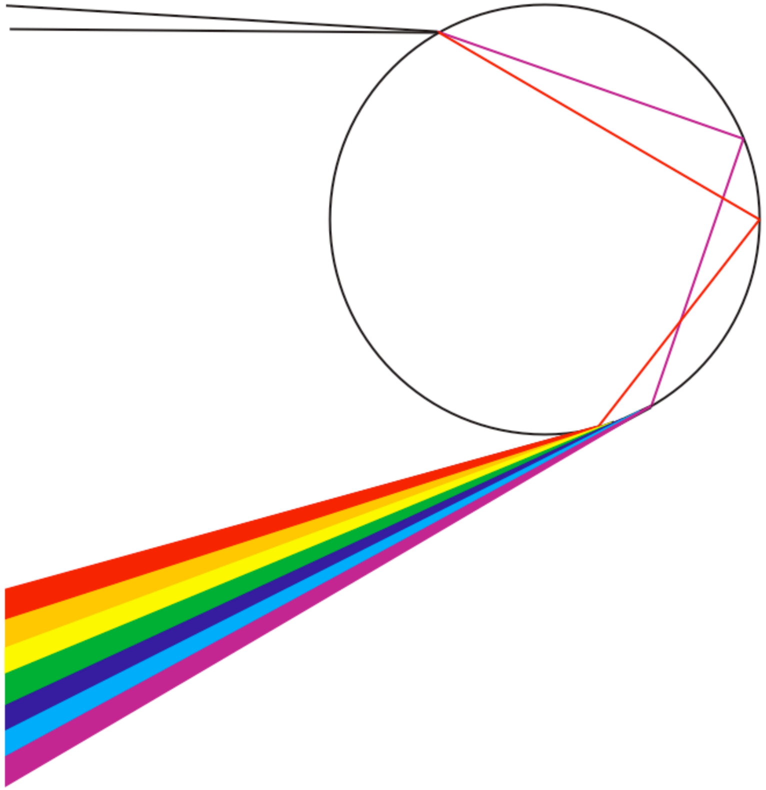
1. Introduction through History
2. Fourier and Wavelet Representations
3. Wavelets and Approximation Theory
4. Wavelets and Compression
5. Going to Two Dimensions: Non-Separale Constructions
6. Beyond Shift Invariant Subspaces
7. Conclusions and Outlook

Outline

1. Introduction through History
    - From Rainbows to Spectras
    - Signal Representations
    - Approximations
    - Compression
2. Fourier and Wavelet Representations
3. Wavelets and Approximation Theory
4. Wavelets and Compression
5. Going to Two Dimensions: Non-Separable Constructions
6. Beyond Shift Invariant Subspaces
7. Conclusions and Outlook

From Rainbows to Spectras



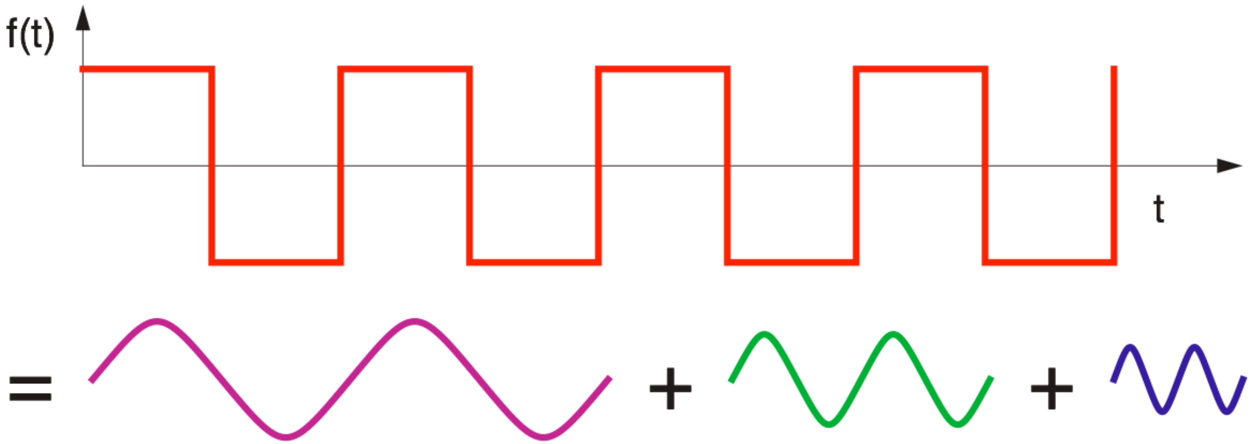


Von Freiberg, 1304: Primary and secondary rainbow

Newton and Goethe

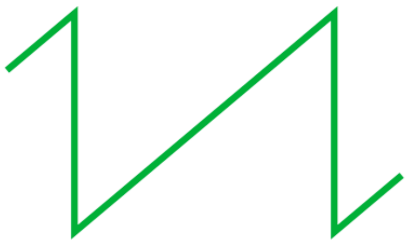
Signal Representations (1/2)

1807: Fourier upsets the French Academy....



Fourier Series: Harmonic series, frequency changes, f_0, $2f_0$, $3f_0$, ...

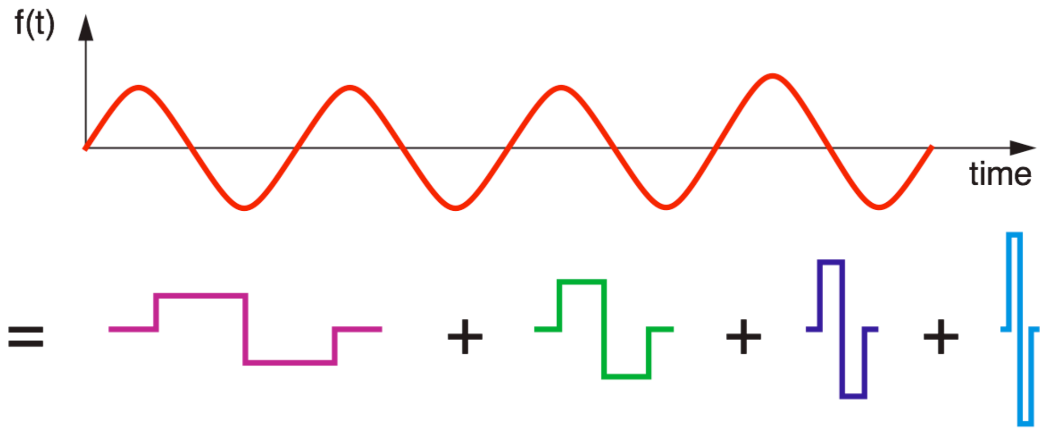
But... 1898: Gibbs' paper



1899: Gibbs' correction

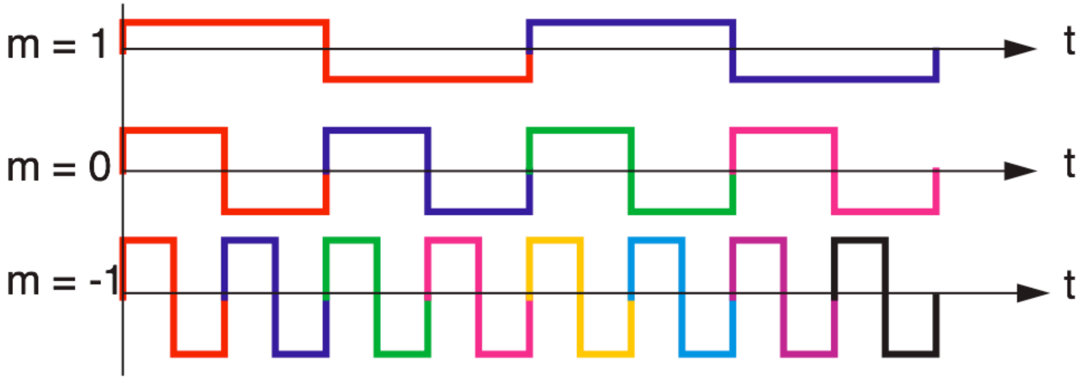
Orthogonality, convergence, complexity

1910: Alfred Haar discovers the Haar wavelet
"dual" to the Fourier construction



Haar series:

- Scale changes S_0, $2S_0$, $4S_0$, $8S_0$ ...
- orthogonality

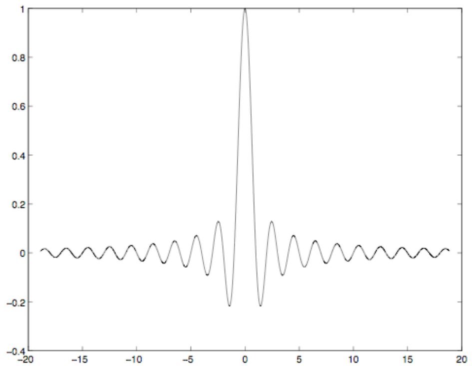


Theorem 1 (Shannon-48, Whittaker-35, Nyquist-28, Gabor-46)

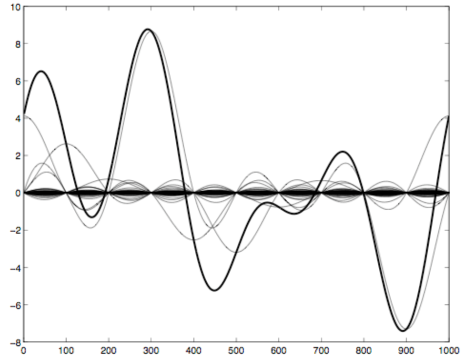
If a function f(t) contains no frequencies higher than W cps, it is completely determined by giving its ordinates at a series of points spaced 1/(2W) seconds apart.
[if approx. T long, W wide, 2TW numbers specify the function]

It is a representation theorem:

- $\{sinc(t-n)\}_{n\ in\ Z}$, is an orthogonal basis for BL$[-\pi,\pi]$
- f(t) in BL$[-\pi,\pi]$ can be written as $f(t)=\sum_n f(n)\cdot sinc(t-n)$



… slow…!



Note:

- Shannon-BW, BL sufficient, not necessary.
- many variations, non-uniform etc
- Kotelnikov-33!

Representations, Bases and Frames

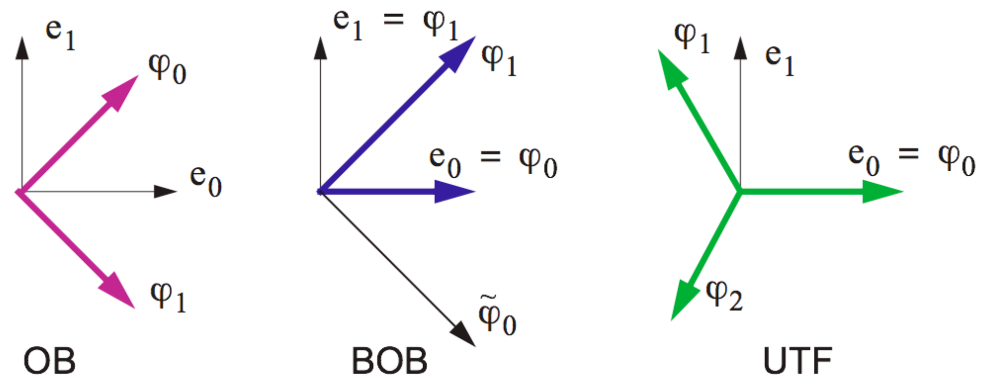
Ingredients:

- as set of vectors, or "atoms", $\{\varphi_n\}$
- an inner product, e.g. $\langle \varphi_n, f \rangle = \int (\varphi_n \cdot f)$
- a series expansion

$$f(t) = \sum_n \langle \varphi_n, f \rangle \cdot \varphi_n(t)$$

Many possibilities:

- orthonormal bases (e.g. Fourier series, wavelet series)
- biorthogonal bases
- overcomplete systems or frames



Note: no transforms, uncountable

Approximations, aproximation…

The linear approximation method

Given an orthonormal basis $\{g_n\}$ for a space S and a signal

$$f = \sum_n \langle f, g_n \rangle \cdot \; g_n$$

the best linear approximation is given by the projection onto a fixed sub-space of size M (independent of f!)

$$\hat{f}_M = \sum_{n \in J_M} \langle f, g_n \rangle \cdot g_n$$

The error (MSE) is thus

$$\varepsilon_M = \| \, f - \hat{f} \, \|^2 = \sum_{n \notin J_M} | \, \langle f, g_n \rangle \, |^2$$

Ex: Truncated Fourier series

project onto first M vectors corresponding to largest
expected inner products, typically LP

The Karhunen-Loeve Transform: The Linear View

Best Linear Approximation in an MSE sense:

Vector processes., i.i.d.:

$$X = [X_0, X_1, \ldots, X_{N-1}]^T \qquad E[X_i] = 0 \qquad E[X \cdot X^T] = R_X$$

Consider linear approximation in a basis

$$\hat{X}_M = \sum_{n=0}^{M-1} \langle X, g_n \rangle \cdot g_n \qquad M < N$$

Then:

$$E[\varepsilon_M] = \sum_{n=M}^{N-1} \langle R_X g_n, g_n \rangle$$

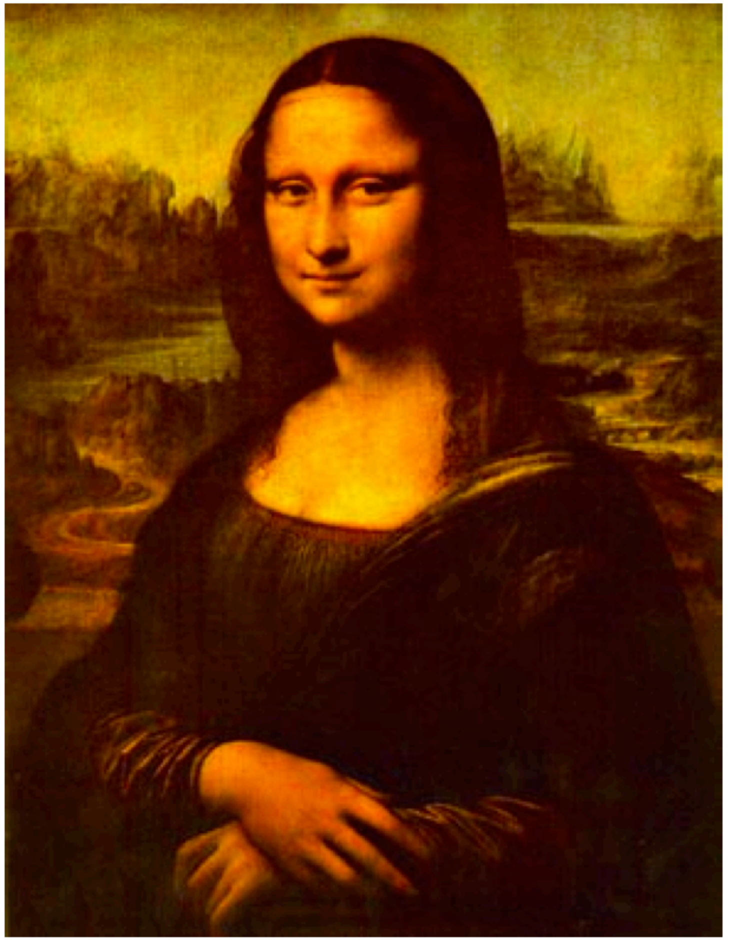
Karhunen-Loeve transform (KLT):

For $0<M<N$, the expected squared error is minimized for the basis $\{g_n\}$ where g_m are the eigenvectors of R_x ordered in order of decreasing eigenvalues.

Proof: eigenvector argument inductively.

Note: Karhunen-47, Loeve-48, Hotelling-33, PCA, KramerM-56, TC

Compression: How many bits for Mona Lisa?



$$\Longleftrightarrow \{0, 1\}$$

A few numbers…

D. Gabor, September 1959 (Editorial IRE)

"... the 20 bits per second which, the psychologists assure us, the human eye is capable of taking in, ..."

Index all pictures ever taken in the history of mankind

- $100 \quad years \cdot 10^{10} \sim 44 \quad bits$

Huffman code Mona Lisa index

- A few bits (Lena Y/N?, Mona Lisa…), what about R(D)….

Search the Web!

- http://www.google.com, 5-50 billion images online, or 33-36 bits

JPEG

- 186K… There is plenty of room at the bottom!
- JPEG2000 takes a few less, thanks to wavelets…

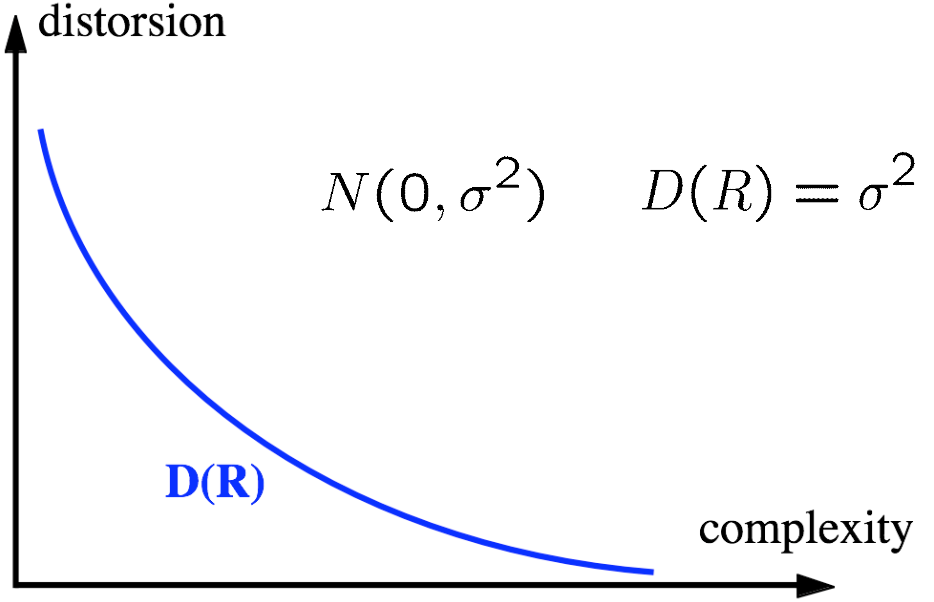
Note: $2^{(256 \times 256 \times 8)}$ possible images (D.Field)

Homework in Cover-Thomas, Kolmogorov, MDL, Occam etc

Source Coding: some background

Exchanging description complexity for distortion:

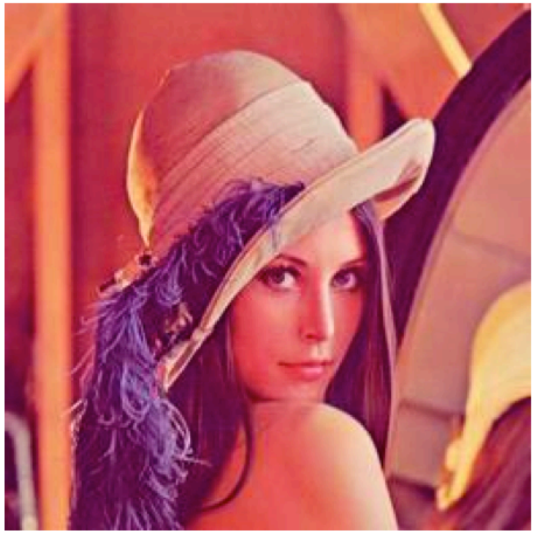
- rate-distortion theory [Shannon:58, Berger:71]
- known in few cases...like i.i.d. Gaussians (but tight: no better way!)



$N(0, \sigma^2)$ $\quad D(R) = \sigma^2 \cdot 2^{-2R}$ $\quad$ or -6dB/bit

- typically: difficult, simple models, high complexity (e.g. VQ)
- high rate results, low rate often unknown

New image coding standard … JPEG 2000



Original Lena Image (256 x 256 Pixels, 24-Bit RGB)

JPEG Compressed (Compression Ratio 43:1)

JPEG2000 Compressed (Compression Ratio 43:1)

From the comparison,

- JPEG fails above 40:1 compression
- JPEG2000 survives

Note: images courtesy of www.dspworx.com

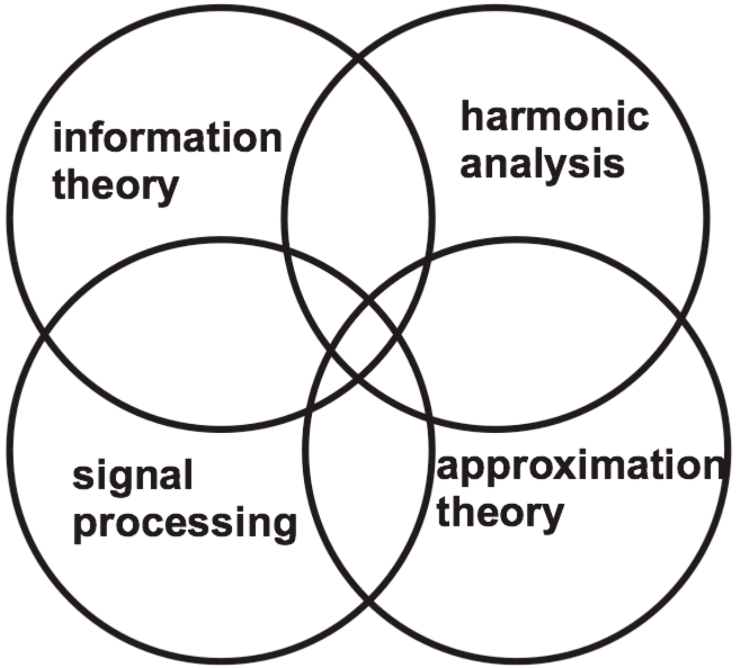
Representation, Approximation and Compression: Why does it matter anyway?

Parsimonious or sparse representation of information is key in

- storage and transmission
- indexing, searching, classification, watermarking
- denoising, enhancing, resolution change

But: it is also a fundamental question in

- information theory
- signal/image processing
- approximation theory
- vision research



Successes of wavelets in image processing:

- compression (JPEG2000)
- denoising
- enhancement
- classification

Thesis: Wavelet models play an important role

Antithesis: Wavelets are just another fad!

Outline

1. Introduction through History
2. Fourier and Wavelet Representations
    - Fourier and Local Fourier Transforms
    - Wavelet Transforms
    - Piecewise Smooth Signal Representations
3. Wavelets and Approximation Theory
4. Wavelets and Compression
5. Going to Two Dimensions: Non-Separable Constructions
6. Beyond Shift Invariant Subspaces
7. Conclusions and Outlook

Fourier and Wavelet Representations: Spaces

Norms:

$$\|x\|_p = \left(\sum_n |x[n]|^p\right)^{\frac{1}{p}} \qquad \|f\|_p = \left(\int_{-\infty}^{\infty} |f(t)|^p \, dt\right)^{\frac{1}{p}}$$

Hilbert spaces:

$$l_2(Z) = \{x : (\|x\|_2 < \infty)\} \qquad L_2(R) = \{f : (\|f\|_2 < \infty)\}$$

Inner product:

$$\langle x, y \rangle = \sum_n x^*[n] y[n] \qquad \langle f, g \rangle = \int f^*(t) g(t) dt$$

Orthogonality:

$$x \perp y \Leftrightarrow \langle x, y \rangle = 0$$

Banach spaces:

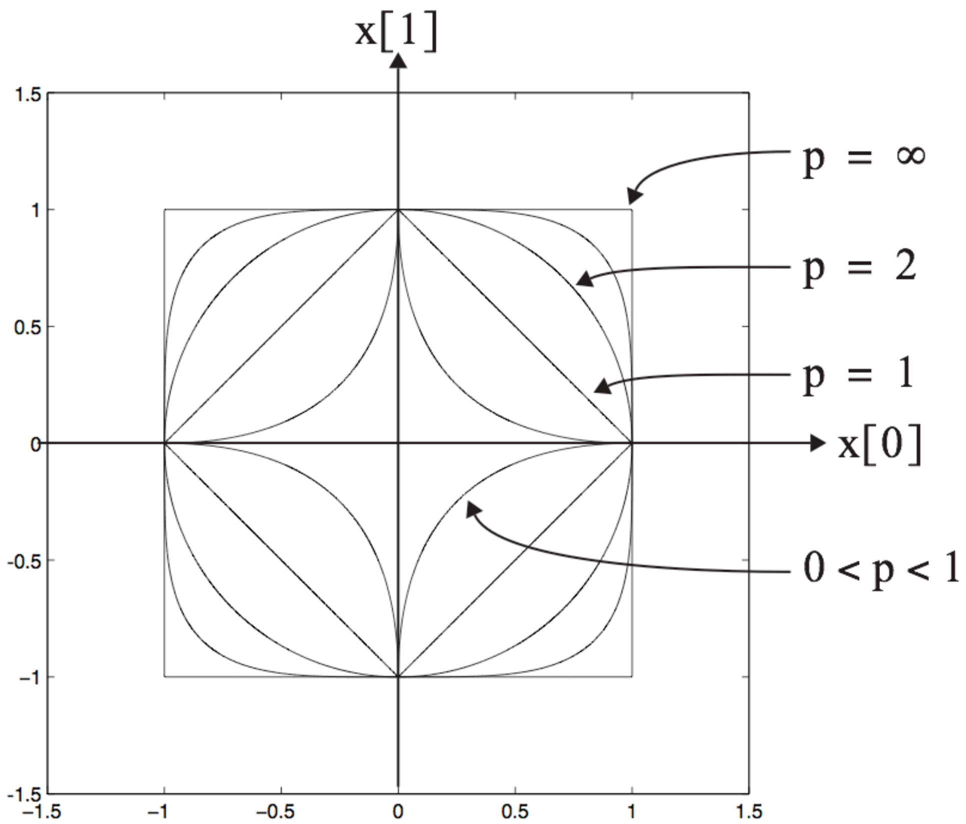
x, f s.t. $\|x\|_p, \|f\|_p < \infty$ $\quad$ p general $\quad$ $p = 1$ and $p = \infty$ of particular interest

C^P spaces: p-times diff. with bounded derivatives -> Taylor expansions

Holder/Lipschitz α : locally α smooth (non-integer)

Example

consider $x \in R^2$ and $||x||_p = 1$



p < 1: quasi norm, p -> 0: sparsity measure

Orthonormal Series Expansion

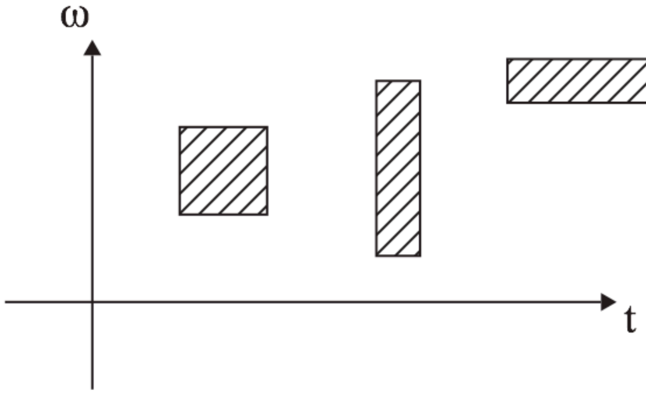
$$f = \sum_{n \in Z} \alpha_n \varphi_n \qquad \alpha_n = \langle \varphi_n, f \rangle \qquad \langle \varphi_n, \varphi_m \rangle = \delta_{n-m} \qquad \|f\|_2 = \|\alpha\|_2$$

Time-Frequency Analysis and Uncertainty Principle

$$f(t) \leftrightarrow F(\omega) \qquad \Delta_t^2 = \int t^2 \,|\, f(t) \,|\, dt \qquad \Delta_\omega^2 = \int \omega^2 \,|\, F(\omega) \,|\, d\omega$$

Then

$$\boxed{\Delta_t^2 \cdot \Delta_\omega^2 \geq \frac{\pi}{2}}$$



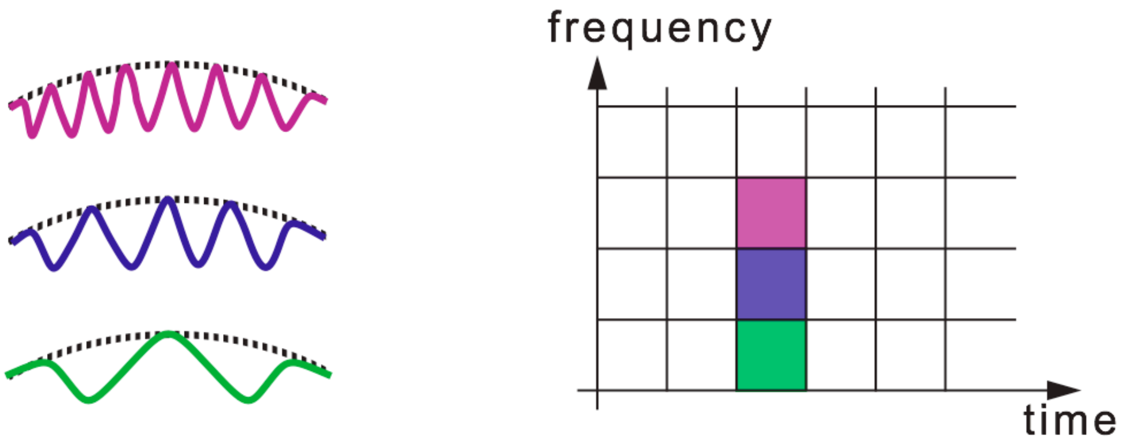
Not arbitrarily sharp in time and frequency!

Local Fourier Basis?

The Gabor or Short-time Fourier Transform

$$\varphi_{m,n}(t) = w(t - nT)e^{-jm\omega_0(t-nT)}$$

Time-frequency atoms localized at $(nT, m\omega_0)$



When T, ω_0 "small enough"

$$f(t) \approx c \cdot F_{m,n}\varphi_{m,n}(t) \quad \text{where} \quad F_{m,n} = \langle \varphi_{m,n}, f \rangle$$

Example: Spectrogram

The Bad News…

Balian-Low Theorem

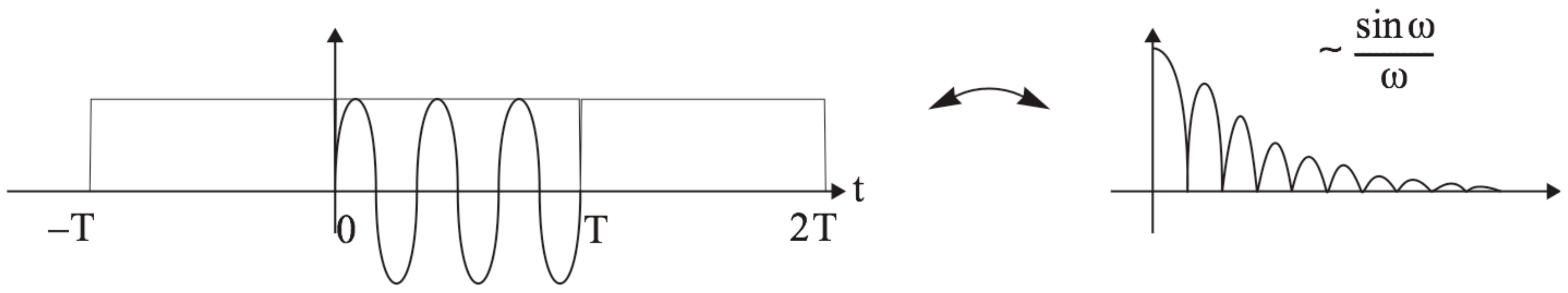
$\varphi_{m,n}$ is a short-time Fourier frame with critical sampling $(T\omega_0 = 2\pi)$

then either

$$\boxed{\Delta_t^2 = \infty \quad \text{or} \quad \Delta_\omega^2 = \infty}$$

or: there is no good local orthogonal Fourier basis!

Example of a basis: block based Fourier series



Note: consequence of BL Theorem on OFDM, RIAA

The Good News!

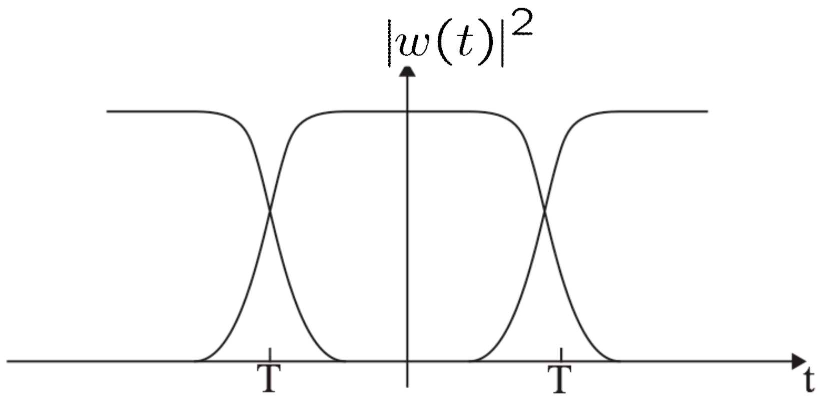
There exist good local cosine bases.

Replace complex modulation ($e^{jm\omega_0 t}$) by appropriate cosine modulation

$$\varphi_{m,n}(t) = w(t-nT)\cos\left(\frac{\pi}{2}\left(m+\frac{1}{2}\right)\left(t-nT+\frac{T}{2}\right)\right)$$

where w(t) is a power complementary window

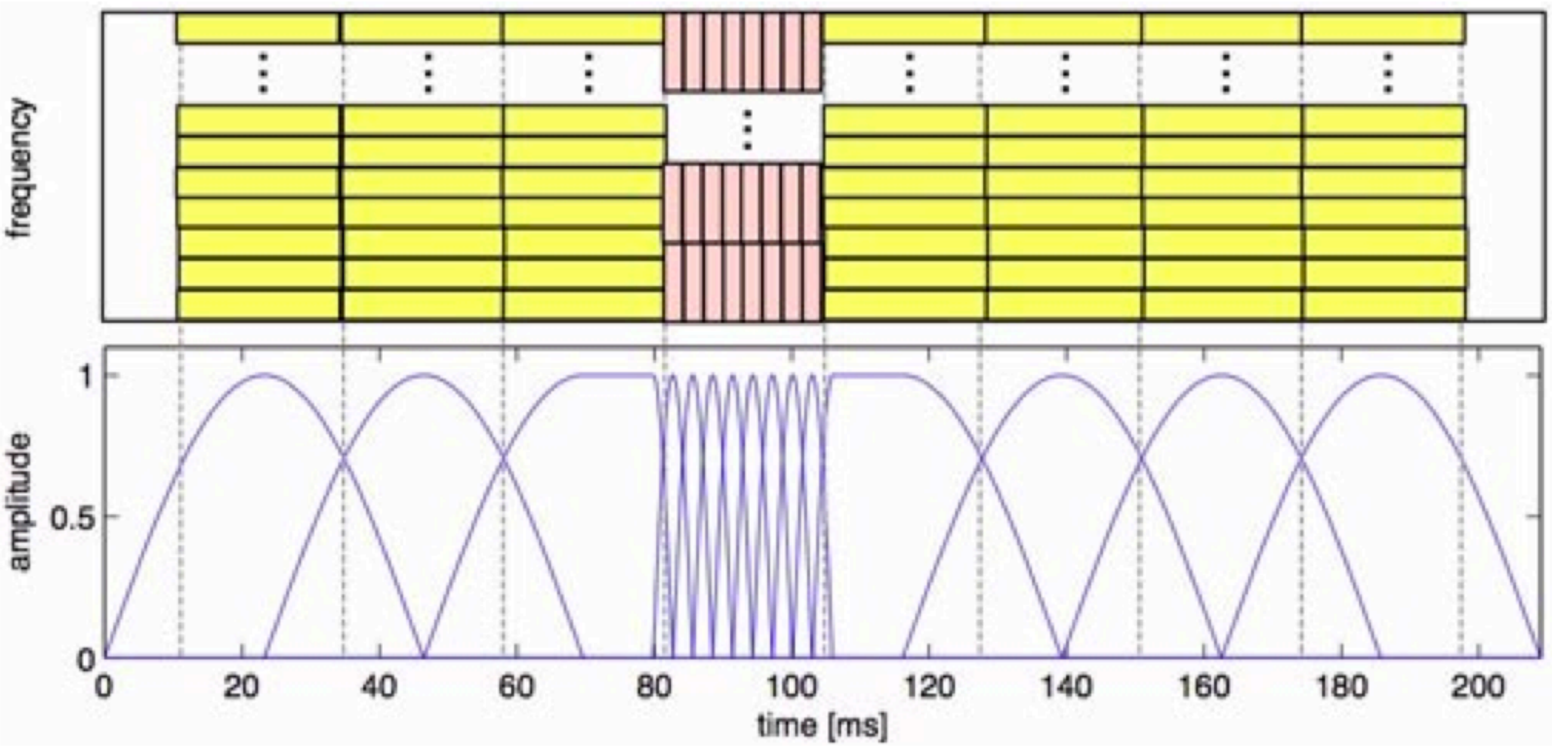
$$\sum_n |W(t-nT)| = 1$$



Result: MP3!

Many generalisations…

Example of time-frequency tiling, state of the art audio coder



In this example, it switches from 1024 channels down to 128, makes for pretty crisp attacks!

It also makes the RIAA nervous….

Another Good News!

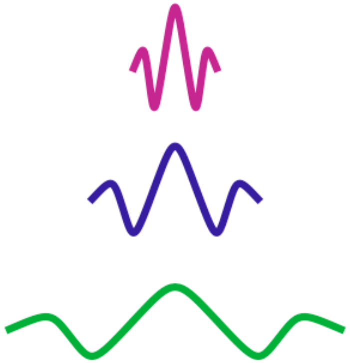
Replace (shift, modulation)

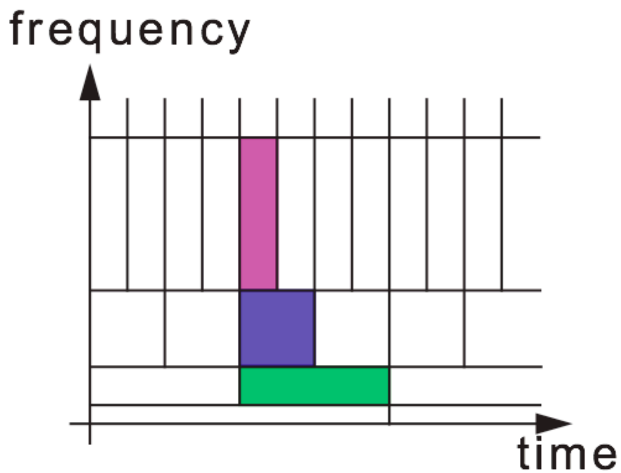
by (shift, scale)

or

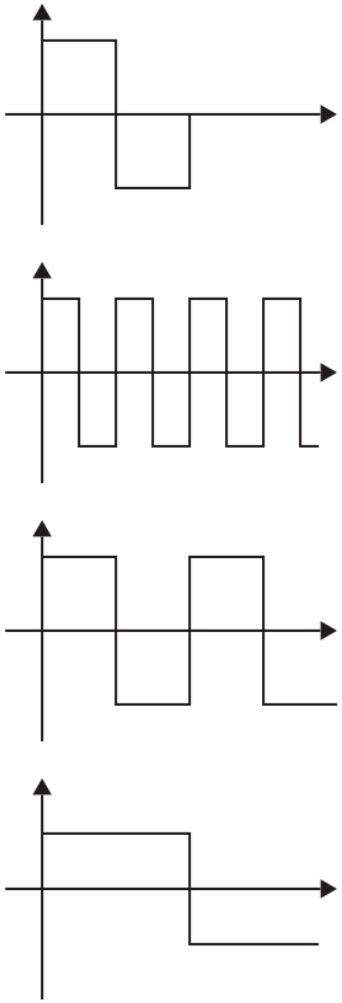
$$\psi_{m,n}(t) = 2^{-\frac{m}{2}}\psi\left(\frac{t-2^m n}{2^m}\right) \qquad n, m \in Z$$

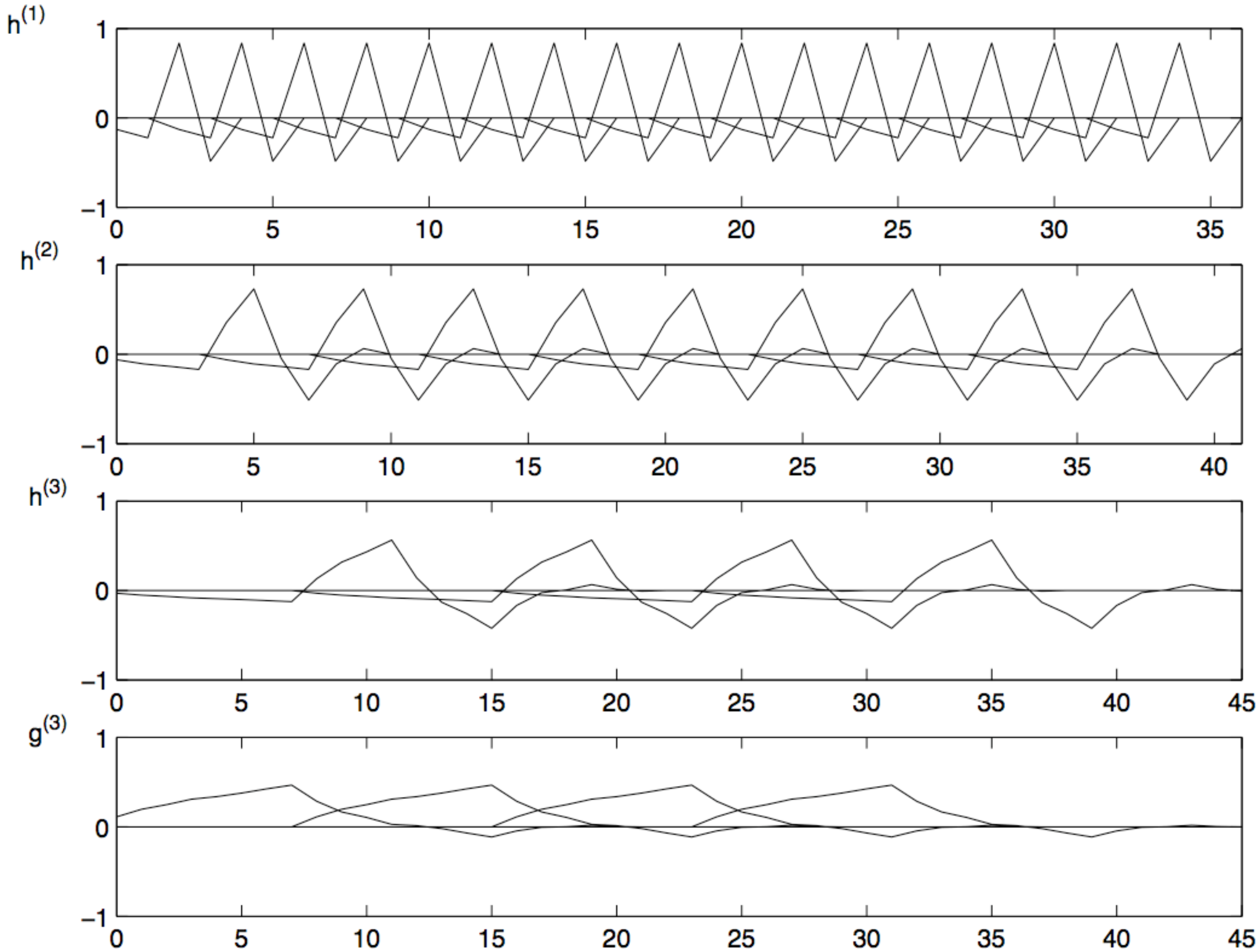
then there exist “good” localized orthonormal bases, or wavelet bases





Examples of bases



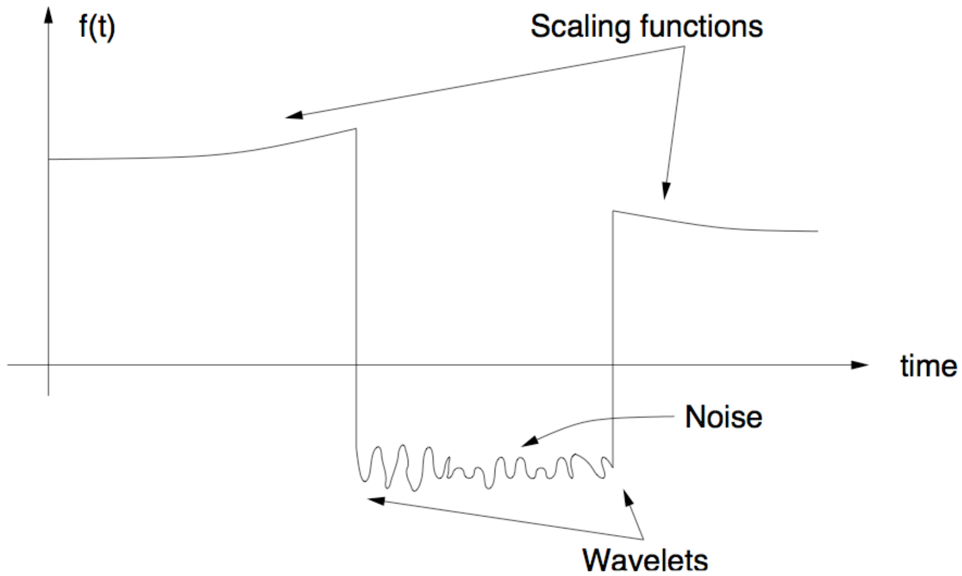


Haar

Daubechies, D_2

Wavelets and representation of piecewise smooth functions

Goal: efficient representation of signals like



where:

- Wavelet act as singularity detectors
- Scaling functions catch smooth parts
- “Noise” is circularly symmetric

Note: Fourier gets all Gibbs-ed up!

Key characteristics of wavelets and scaling functions (1/3)

Daubechies-88, Wavelets from filter banks, ortho-LP with N zeros at π ,

$$G(z) = (1 + z^{-1})^N \cdot R(z)$$

Scaling function:

$$\varphi(\omega) = \prod_{i=1}^{\infty} G\left(e^{j\left(\frac{\omega}{2^i}\right)}\right)$$

Orthonormal wavelet family:

$$\psi_{m,n}(t) = 2^{-\frac{m}{2}} \psi(2^{-m} t - n)$$

Scaling function and approximations

- Scaling function $\varphi(t)$ spans polynomials up to degree N-1

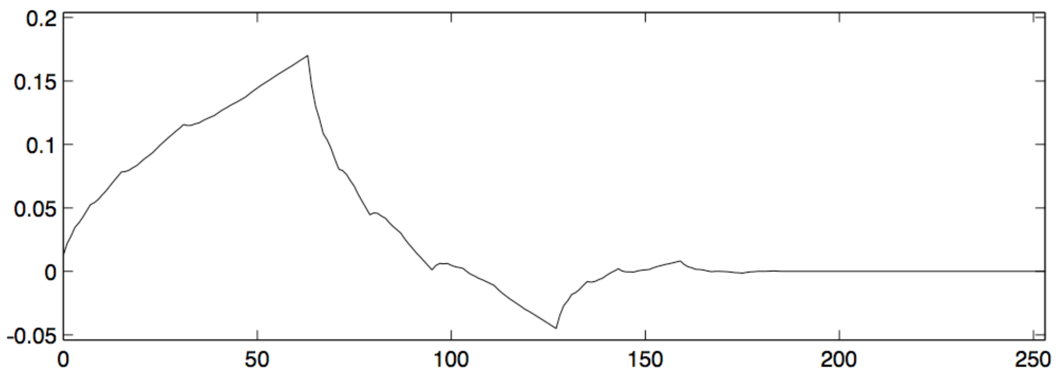
$$\sum_n c_n \cdot \varphi(t-n) = t^k \qquad k = 0, 1, ..., N-1$$

- Strang-Fix theorem: if $\varphi(\omega)$ has N zeros at multiples of 2π (but the origin), then $\{\varphi(t-n)\}_{n \in Z}$ spans polynomials up to degree N-1
- Two scale equation: $$\varphi(t) = \frac{1}{\sqrt{2}} \cdot \sum_n g_n \cdot \varphi(2t - n)$$
- Smoothness: follows from N, $\alpha = 0,203\ N$

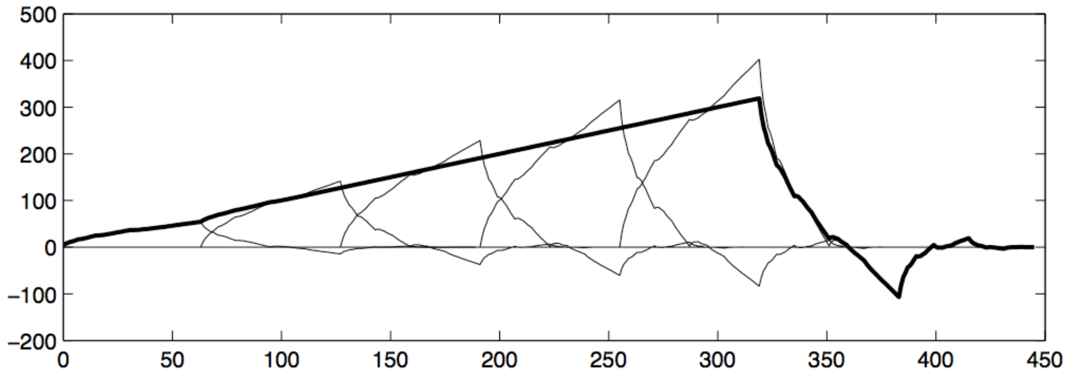
Key characteristics of wavelets and scaling functions (2/3)

Lowpass filters and scaling functions reproduce polynomials

- Iterate of Daubechies L=4 lowpass filter reproduces linear ramp



scaling function



linear ramp

Scaling functions catch "trends" in signals

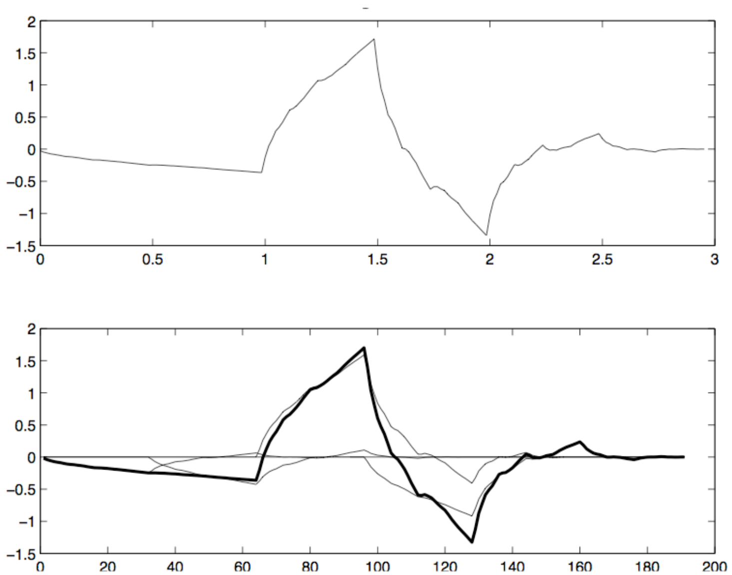
Key characteristics of wavelets and scaling functions (3/3)

Wavelet approximations

- wavelet ψ has N zeros moments, kills polynomials up to degree N-1
- wavelet of length L=2N-1, or 2N-1 coeffs influenced by singularity at each scale, wavelet are singularity detectors,
- wavelet coefficients of smooth functions decays fast,
  e.g. f in c^P, m << 0

$$\langle \psi_{m,n}, f \rangle = c 2^{m(p-\frac{1}{2})}$$

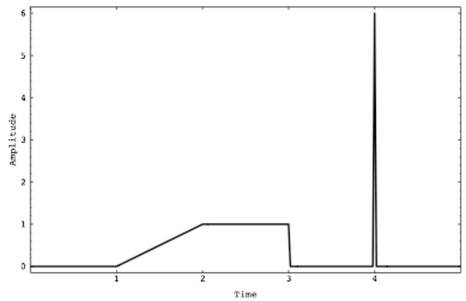
Note: all this is in 1 dimension only, 2D is another story…

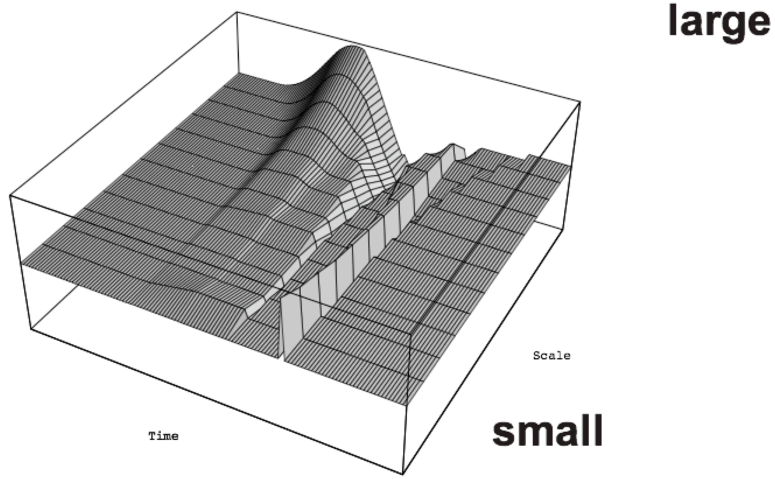


How about singularities?

If we have a singularity of order n at the origin (0: Dirac, 1: Heaviside,…), the CWT transform behaves as

$$X(a,0) = c_n \cdot a^{(n-\frac{1}{2})}$$



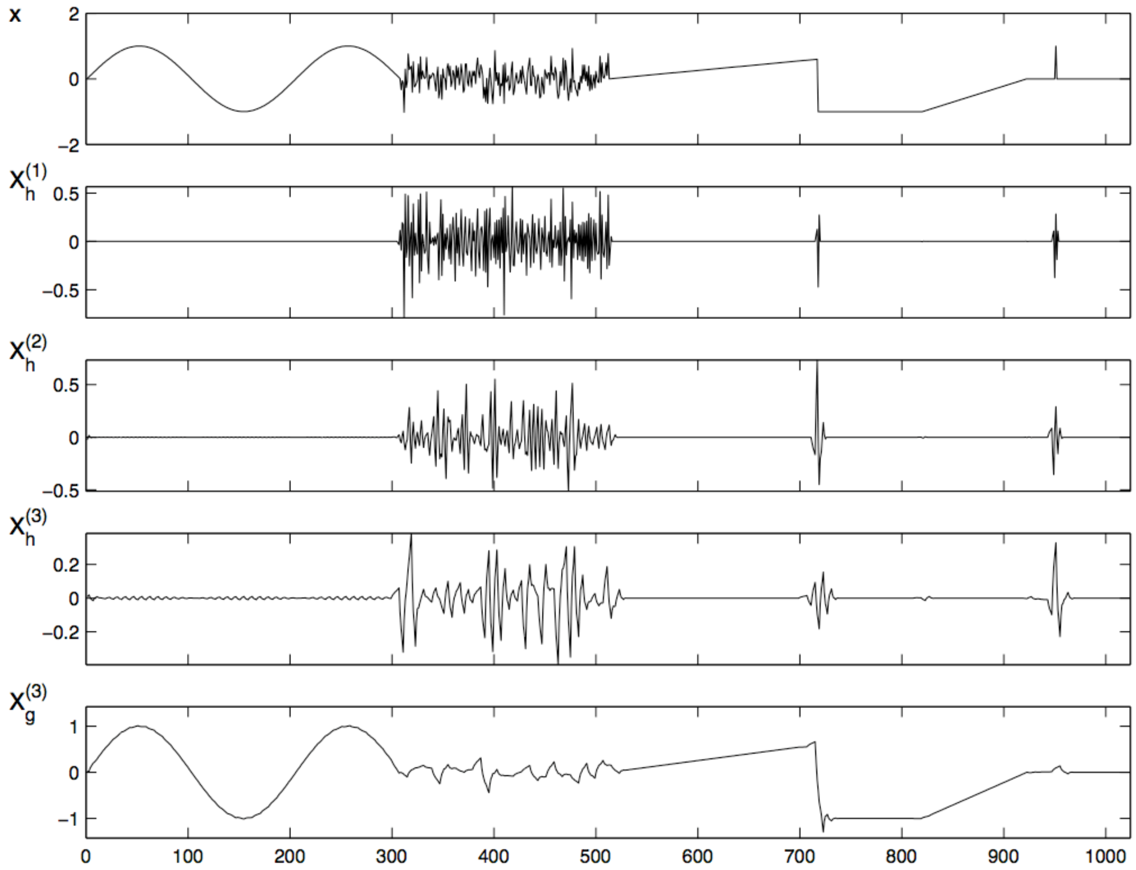


In the orthogonal wavelet series: same behavior, but only L=2N-1 coefficients influenced at each scale!

- E.g. Dirac/Heaviside: behavior as $2^{-\frac{m}{2}}$ and $2^{\frac{m}{2}}, \;\; m << 0$

Wavelets catch and characterize singularities!

Thus: a piecewise smooth signal expands as:



- phase changes randomize signs, but not decay
- a singularity influence only L wavelets at each scale
- wavelet coefficients decay fast

Outline

1. Introduction through History
2. Fourier and Wavelet Representations
3. Wavelets and Approximation Theory
    - Non-linear approximation
    - Fourier versus wavelet, LA versus NLA
4. Wavelets and Compression
5. Going to Two Dimensions: Non-Separable Constructions
6. Beyond Shift Invariant Subspaces
7. Conclusions and Outlook

From linear to non-linear approximation theory

The non-linear approximation method

Given an orthonormal basis $\{g_n\}$ for a space S and a signal

$$f = \sum_n \langle f, g_n \rangle \cdot g_n$$

the best nonlinear approximation is given by the projection onto an adapted subspace of size M (dependent on f!)

$$\tilde{f}_M = \sum_{n \in I_M} \langle f, g_n \rangle \cdot g_n$$

$$I_M : \quad |\langle f, g_n \rangle|_{n \in I_M} \geq |\langle f, g_m \rangle|_{m \notin I_M}$$ set of M largest $\langle\ ,\ \rangle$

The error (MSE) is thus $\tilde{\varepsilon}_M = \| f - \tilde{f} \|^2 = \sum_{n \notin I_M} |\langle f, g_n \rangle|^2$

and $\tilde{\varepsilon}_M \leq \varepsilon_M$

Difference: take the first M coeffs (linear) or

take the largest M coeffs (non-linear)

From linear to non-linear approximation theory

Nonlinear approximation

- This is a simple but nonlinear scheme
- Clearly, if $A_M(.)$ is the NL approximation scheme:

$$A_M(x) + A_M(y) \neq A_M(x+y)$$

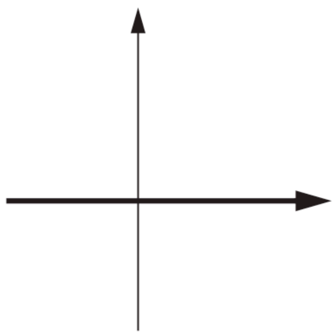
This could be called "adaptive subspace fitting"

From a compression point of view, you "pay" for the adaptivity

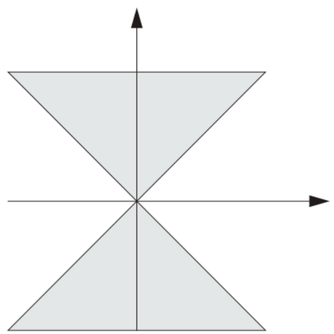
- in general, this will cost

$$\log\left(\begin{bmatrix} N \\ k \end{bmatrix}\right) \text{ bits}$$

These cannot be spent on coefficient representation anymore



LA: pick a subspace a priori

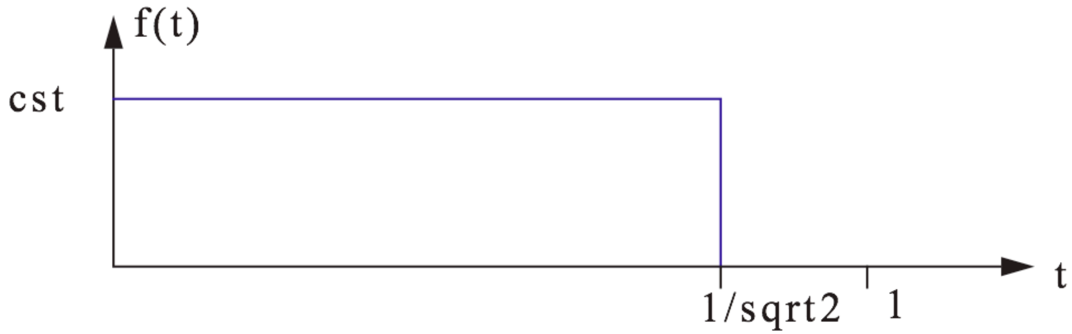


NLA: pick after seeing the data

Non-Linear Approximation Example

Nonlinear approximation power depends on basis

Example:



Two different bases for $[0, 1]$:

- Fourier series $\{e^{j2\pi kt}\}_{k\in F}$
- Wavelet series: Haar wavelets

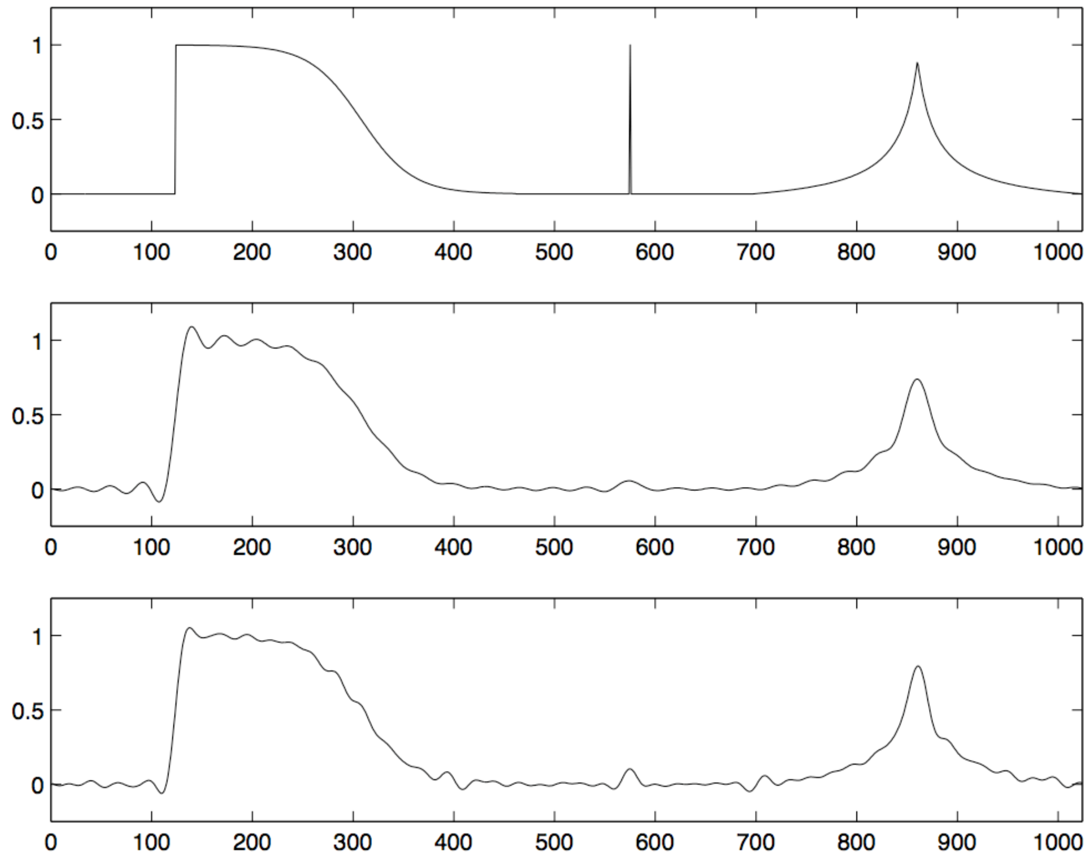
Linear approximation in Fourier or wavelet bases $\varepsilon_M \sim \frac{1}{M}$

Nonlinear approximation in a Fourier basis $\tilde{\varepsilon}_M \sim \frac{1}{M}$

Nonlinear approximation in a wavelet basis $\tilde{\varepsilon}_M \sim \frac{1}{2^M}$

Non-linear Approximation Example

Fourier basis: N=1024, M=64, linear versus nonlinear



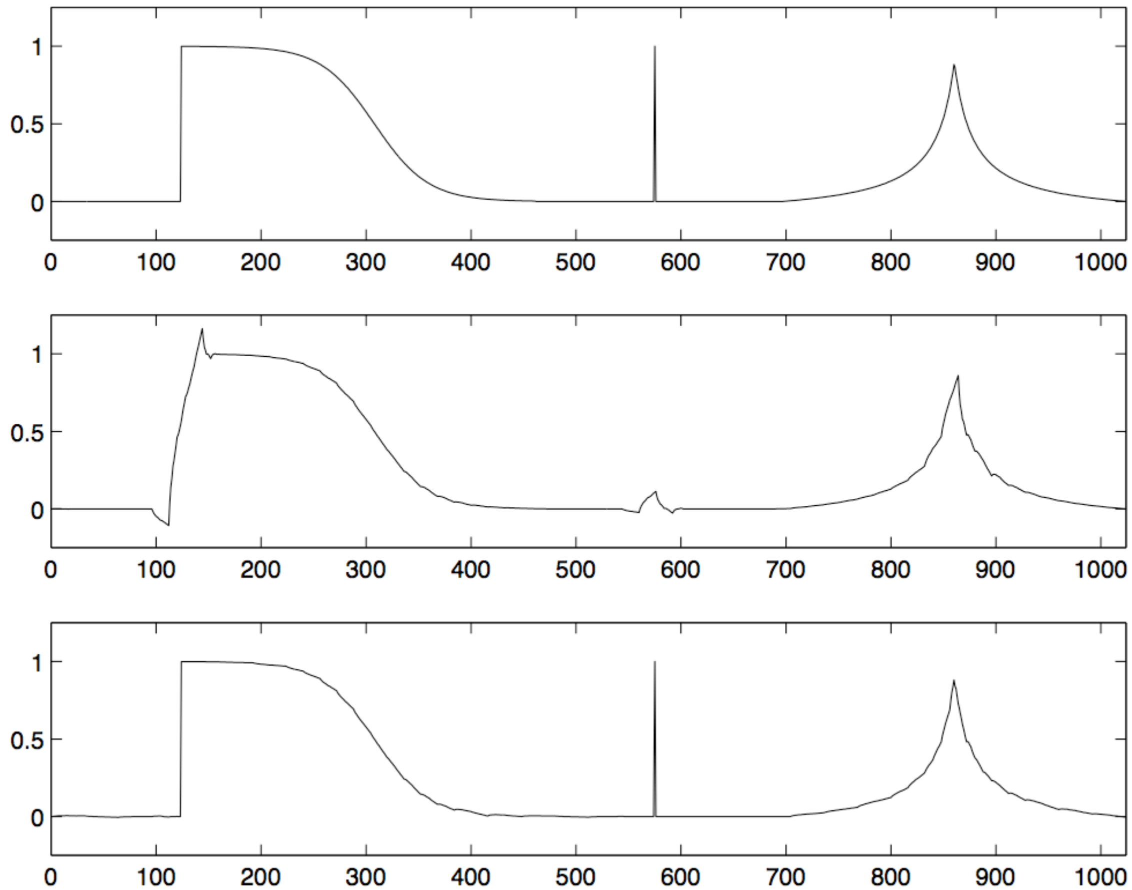
D= 2.7

D= 2.4

- Nonlinear approximation is not necessarily much better!

Non-linear Approximation Example

Wavelet basis: N=1024, M=64, J=6, linear versus nonlinear

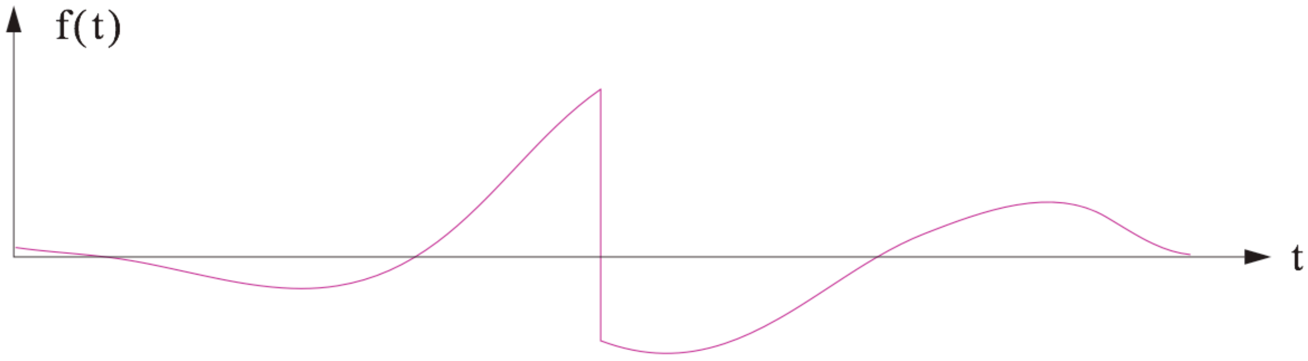


D= 3.5

D= 0.01

- Nonlinear approximation is vastly superior!

Nonlinear approximation theory and wavelets



Approximation results for piecewise smooth fcts

- between discontinuities,
  behavior by Sobolev or Besov regularity
- k derivatives $\Rightarrow$ coeffs $\sim 2^{m(k-\frac{1}{2})}$ when $m \ll 0$
- Besov spaces can be defined with wavelets bases. If

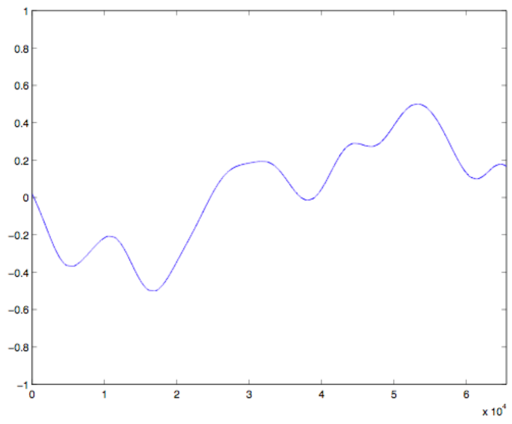
$$\| f \|_{G,p} = \left(\sum |\langle f, g_n \rangle|^p \right)^{\frac{1}{p}} < \infty \qquad 0 < p < 2$$

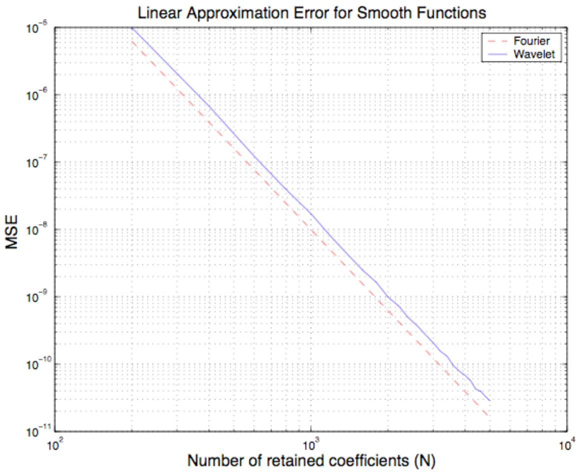
then [DeVoreJL92]:

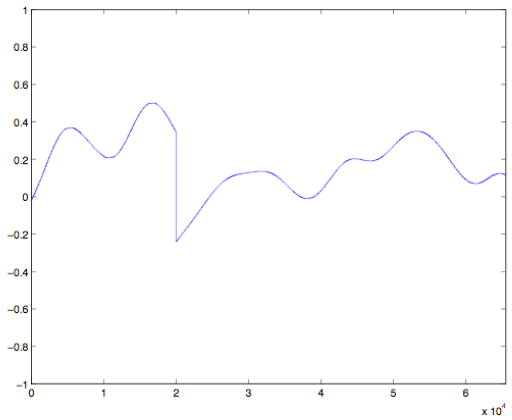
$$\tilde{\varepsilon}_M = o(M^{1-\frac{2}{p}})$$

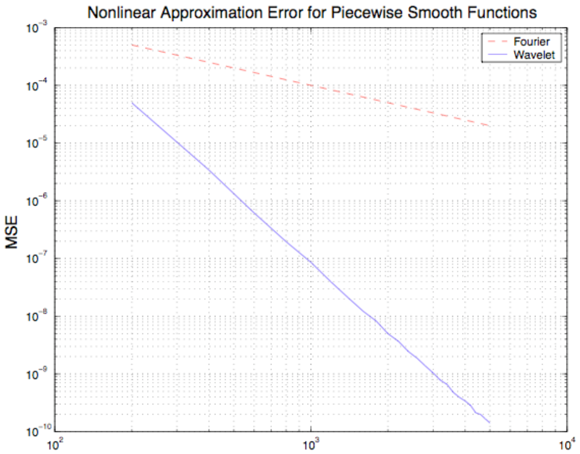
Smooth versus piecewise smooth functions:

It depends on the basis and on the approximation method









s=2, N=2^16, D_3, 6 levels

Outline

1. Introduction through History
2. Fourier and Wavelet Representations
3. Wavelets and Approximation Theory
4. Wavelets and Compression
    - A small but instructive example
    - Piecewise polynomials and D(R)
    - Piecewise smooth and D(R)
    - Improved wavelet schemes
5. Going to Two Dimensions: Non-Separable Constructions
6. Beyond Shift Invariant Subspaces
7. Conclusions and Outlook

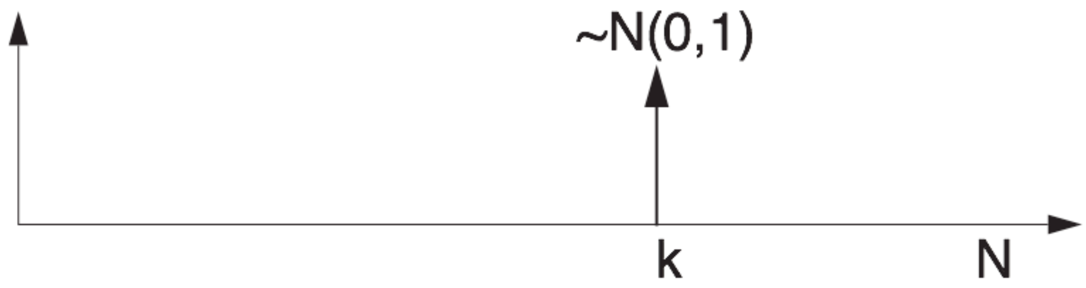
Wavelets and Compression

Compression is just one bit trickier than approximation…

A small but instructive example:

Assume

- $x[n] = \alpha\delta[n-k]$, signal is of length N, k is U[0, N-1] and α is $N(0,1)$
- This is a Gaussian RV at location k



- Note: $R_x = I$!

Linear approximation: $\varepsilon_M = \dfrac{1}{M}$

Non-linear approximation, M > 0: $\tilde{\varepsilon}_M = 0$

Given budget R for block of size N:

1. **Linear approximation and KLT: equal distribution of R/N bits**

$$D(R) = c \cdot \sigma^2 \cdot 2^{-2(R/N)}$$

This is the optimal linear approximation and compression!

2. **Rate-distortion analysis [Weidmann:99]**

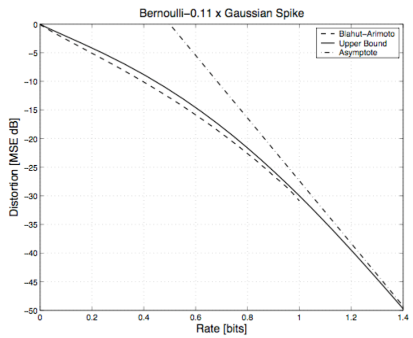
High rate cases:

- Obvious scheme: pointer + quantizer

$$D(R) = c \cdot \sigma^2 \cdot 2^{-2(R-logN)}$$

- This is the R(D) behavior for R >> Log N
- Much better than linear approximation

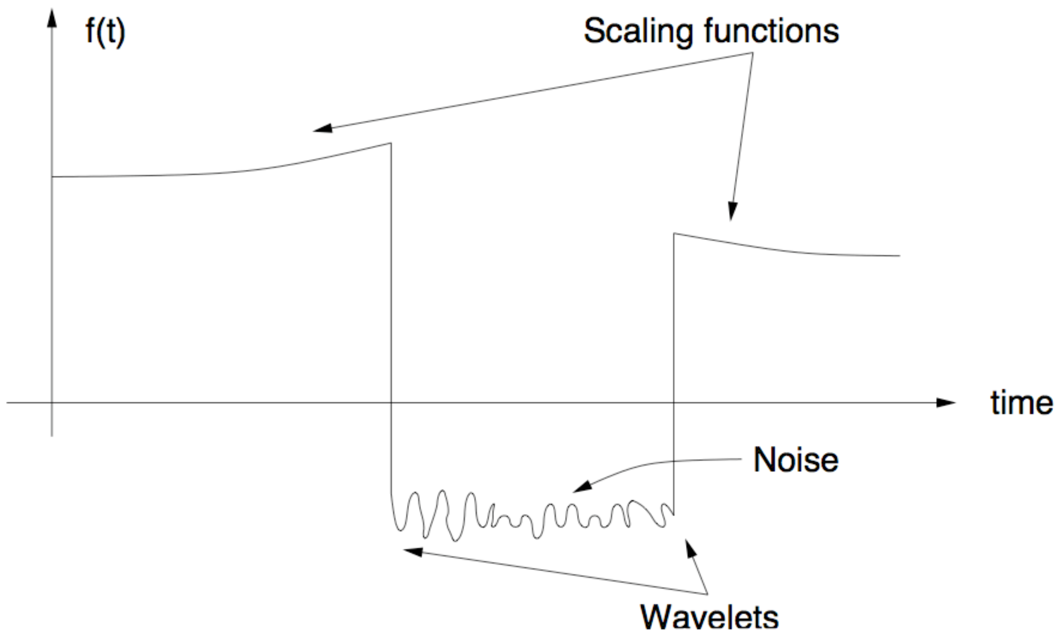
p=0.11



Low rate case:

- Hamming case solved, inc. multiple spikes:
  - there is a linear decay at low rates
- L2 case: upper bounds that beat linear approx.

Piecewise smooth functions: pieces are Lipschitz-α



The following D(R) behavior is reachable [CohenDGO:02]:

$$D(R) = c_1 \cdot R^{-2\alpha} + c_3 \cdot \sqrt{R} \cdot 2^{-c_4 \cdot \sqrt{R}}$$

There are 2 modes:

- $R^{-2\alpha}$ corresponding to the Lipschitz-α pieces
- $\sqrt{R} \cdot 2^{-c \cdot \sqrt{R}}$ corresponding to the discontinuities

Piecewise polynomial, with max degree N

A. **Nonlinear approximation with wavelets having $N+1$ zero moments**

$$D_w(R) = C'_w \cdot (1 + \alpha\sqrt{C_W R}) \cdot 2^{-\sqrt{C_W R}}$$

B. **Oracle-based method**

$$D_p(R) = C'_p \cdot 2^{-(C_p \cdot R)}$$

Thus

- wavelets are a generic but suboptimal scheme
- oracle method asymptotically superior but dependent on the model

Conclusion on compression of piecewise smooth functions:
D(R) behavior has two modes:

$$D(R) = c_1 \cdot R^{-2\alpha} + c_3 \cdot \sqrt{R} \cdot 2^{-c_4 \cdot \sqrt{R}}$$

- 1/polynomial decay: cannot be (substantially) improved
- exponential mode: can be improved, important at low rates

Can we improve wavelet compression?

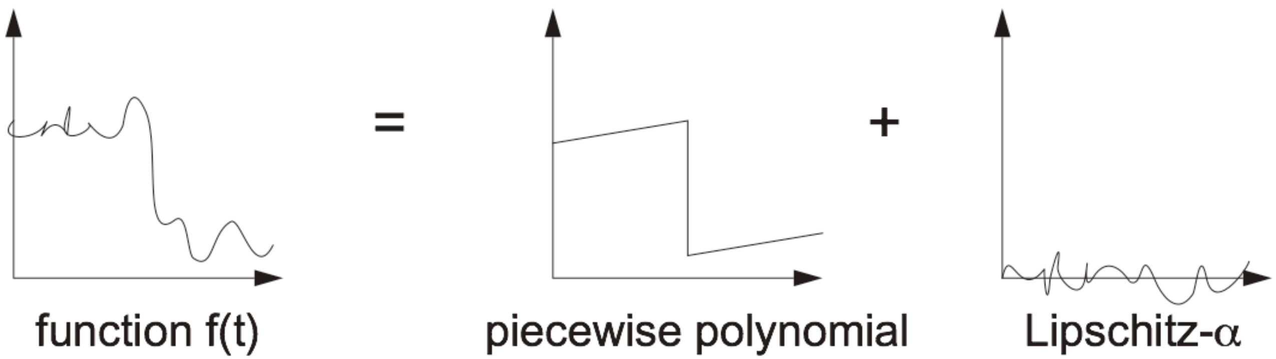
Key: Remove depencies across scales:

- dynamic programming: Viterbi-like algorithm
- tree based algorithms: pruning and joining
- wavelet footprints: wavelet vector quantization

Theorem [DragottiV:03]:

Consider a piecewise smooth signal f(t), where pieces are Lipschitz-α.
There exists a piecewise polynomial p(t) with pieces of maximum degree $\lfloor \alpha \rfloor$
such that the residual $r_\alpha(t) = f(t) - p(t)$
is uniformly Lipschitz-α.

This is a generic split into piecewise polynomial and smooth residual



Footprint Basis and Frames

Suboptimality of wavelets for piecewise polynomials is due to independent coding of dependent wavelet coefficients

$$D_W(R) \sim C \cdot \sqrt{R} \cdot 2^{-\sqrt{R}}$$

Compression with wavelet footprints

Theorem: [DragottiV:03]

Given a bounded piecewise polynomial of deg D with K discontinuities. Then, a footprint based coder achieves

$$D(R) = c_1 \cdot 2^{-(c_2 \cdot R)}$$

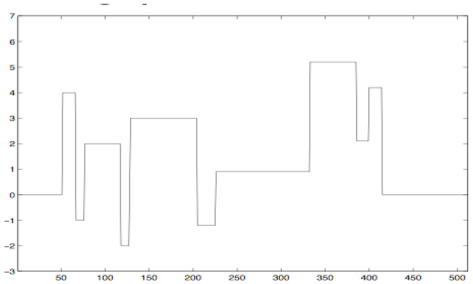
This is a computational effective method to get oracle performance

What is more, the generic split “piecewise smooth” into “uniformly smooth + piecewise polynomial” allows to fix wavelet scenarios, to obtain

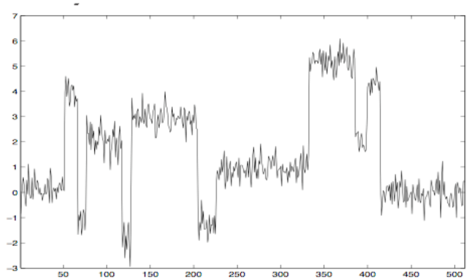
$$D(R) = c_1 \cdot R^{-2\alpha} + c_2 \cdot 2^{-c_3 \cdot R}$$

This can be used for denoising and superresolution

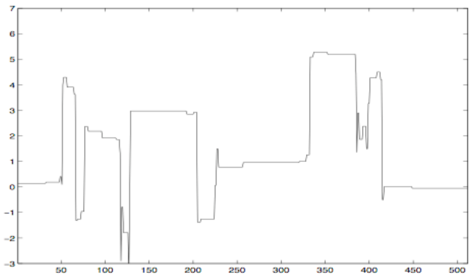
Denoising (use coherence across scale)



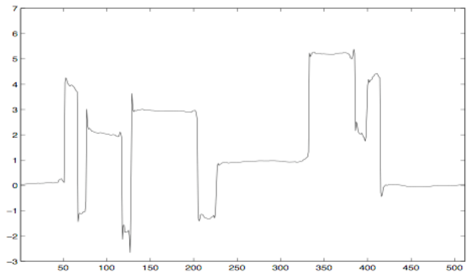
Original signal



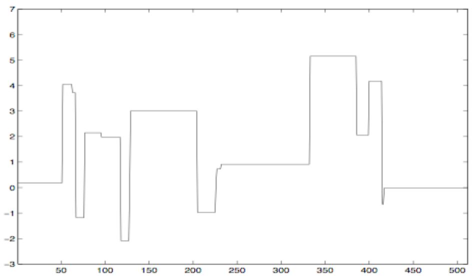
Noisy Signal (SNR=15.62dB)



Hard-Thresholding (SNR=21.3dB)



Cycle-Spinning (SNR=25.4dB)



Denoising with Footprints (SNR=27.2dB)

This is a vector thresholding method adapted to wavelet singularities

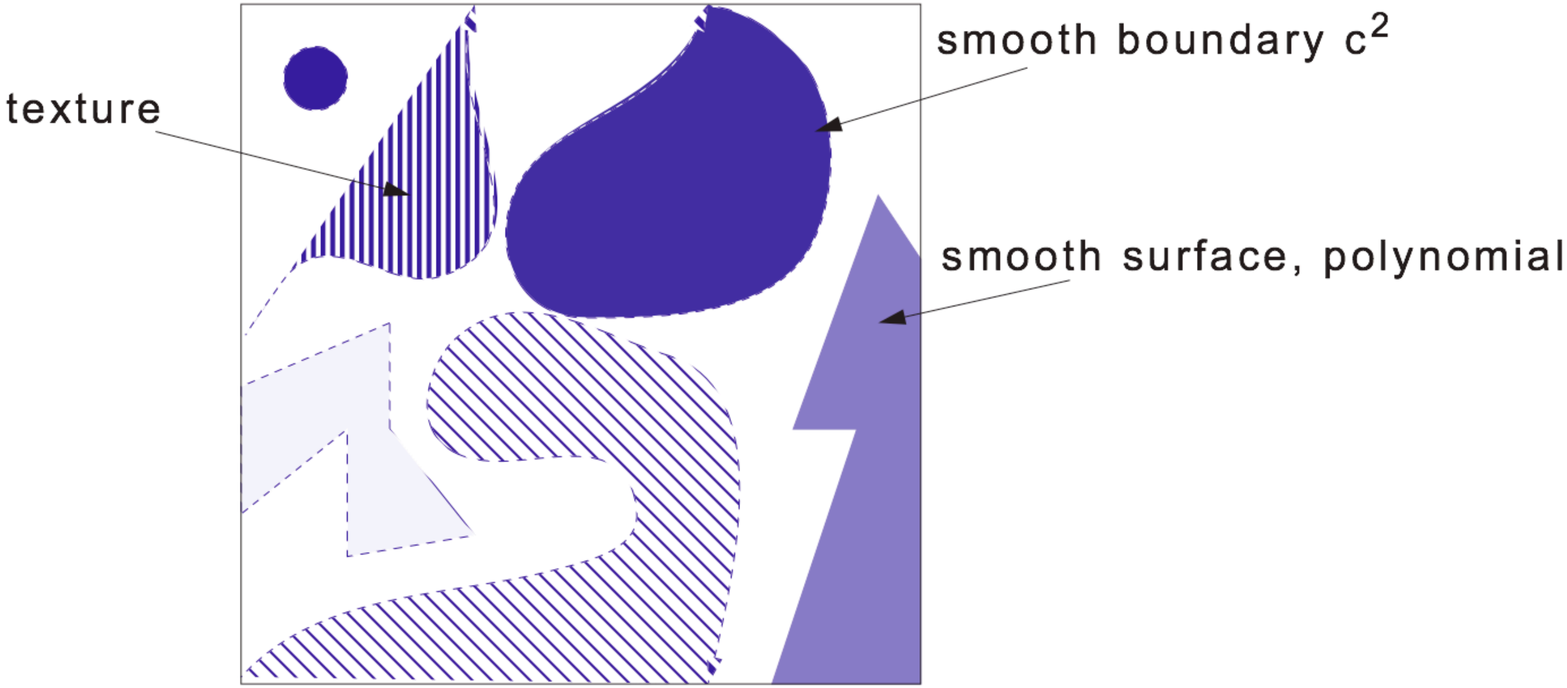
Outline

1. Introduction through History
2. Fourier and Wavelet Representations
3. Wavelets and Approximation Theory
4. Wavelets and Compression
5. Going to Two Dimensions: Non-Separable Constructions
    - The need for truly two-dimensional constructions
    - Tree based methods
    - Non-separable bases and frames
6. Beyond Shift Invariant Subspaces
7. Conclusions and Outlook

Going to Two Dimensions: Non-Separable Constructions

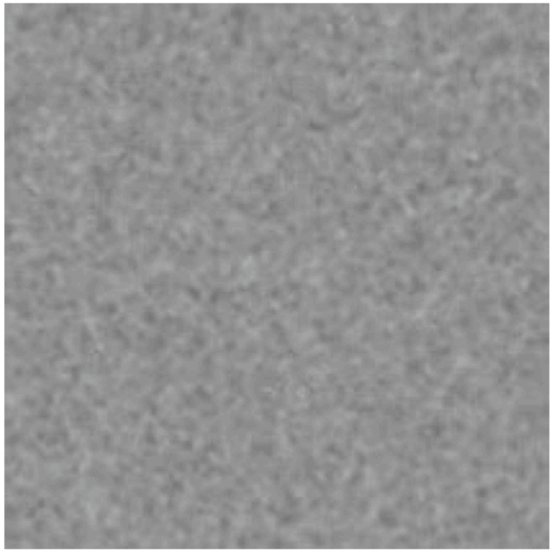
Going to two dimensions requires non-separable bases

Objects in two dimensions we are interested in

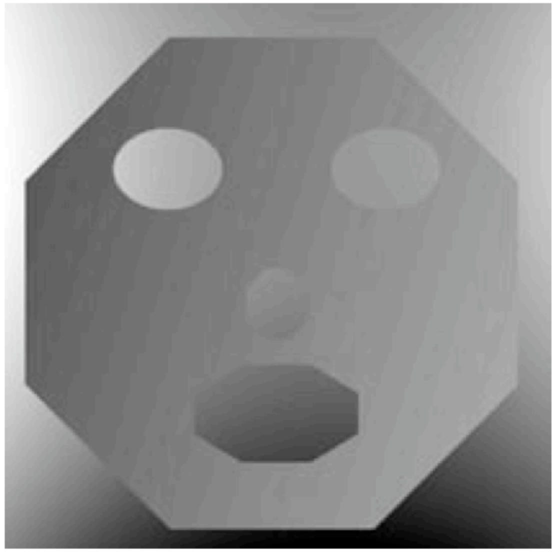


- textures: $D(R) = C_0 \cdot 2^{-2R}$ per pixel
- smooth surfaces: $D(R) = C_1 \cdot 2^{-2R}$ per object!

Models of the world:



Gauss-Markov



Piecewise polynomial

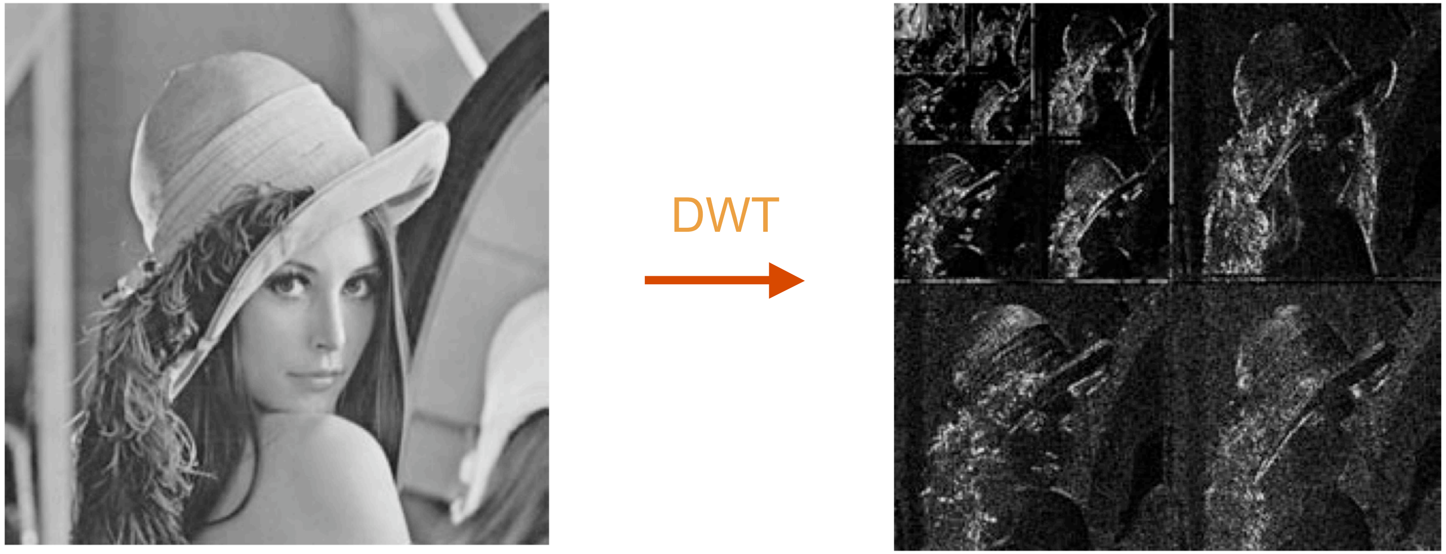
the usual suspect

Many proposed models:

- mathematical difficulties
- one size fits all…
- reality check
- Lena is not PC, but is she BV?

Current approaches to two dimensions….

Mostly separable, direct or tensor products



Fourier and wavelets are both direct product constructions

Wavelets: good for point singularities but what is needed are sparse coding of edge singularities!

- 1D: singularity 0-dimensional (e.g. spike, discontinuity)
- 2D: singularity 1-dimensional (e.g. smooth curve)

Recent work on geometric image processing

Long history: compression, vision, filter banks

Current affairs:

Signal adapted schemes

- Bandelets [LePennec & Mallat]: wavelet expansions centered at discontinuity as well as along smooth edges
- Non-linear tilings [Cohen, Mattei]: adaptive segmentation
- Tree structured approaches [Shukla et al, Baraniuk et al]

Bases and Frames

- Wedgelets [Donoho]: Basic element is a wedge
- Ridgelets [Candes, Donoho]: Basic element is a ridge
- Curvelets [Candes, Donoho]
  Scaling law: width ~length^2
  L(R^2) set up
- Multidirectional pyramids and contourlets [Do et al]
  Discrete-space set-up, l(Z^2)
  Tight frame with small redundancy
  Computational framework

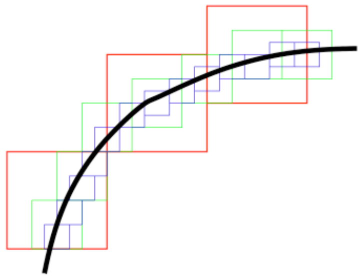
Nonseparable schemes and approximation

Approximation properties:

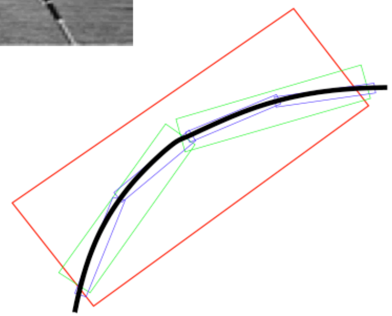
- wavelets good for point singularities
- ridgelets good for ridges
- curvelets good for curves

Consider c^2 boundary between two csts

 # wavelet coeffs $O(2^j)$



 # curvelet coeffs $O(2^{j/2})$



Rate of approximation, M-term NLA in bases, c^2 boundary

- Fourier: $O(M^{-1/2})$
- Wavelets: $O(M^{-1})$
- Curvelets: $O(M^{-2})$

Note: adaptive schemes, Bandelets: $O(M^{-\alpha})$

Multiresolution directional filterbanks and contourlets [M.Do]

Idea: find a direct discrete-space construction that has good approximation properties for smooth functions with smooth boundaries

- directional analysis as in a Radon transform
- multiresolution as in wavelets and pyramids
- computationally easy
- bases or low redundancy frame

Background:

- curvelets [Candes-Donoho] indicate that “good” fixed bases do exist for approximation of piecewise smooth 2D functions
- a frequency-direction relationship indicates a scaling law $d \sim j^{1/2}$
- an effective compression algorithm requires
  - close to a basis (e.g. tight frame with low redundancy)
  - discrete-space set up and computationally efficient

Question:

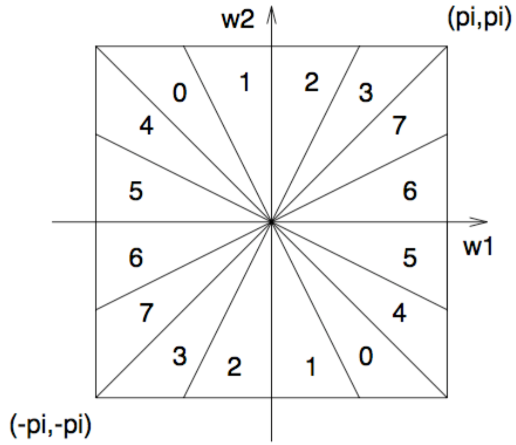
- can we go from $l(Z^2)$ to $L(R^2)$, just like filter banks lead to wavelets?

Answer:

- contourlets!

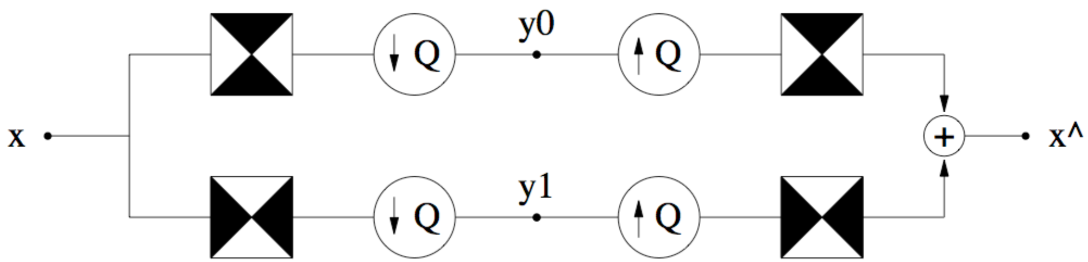
Directional Filter Banks [BambergerS:92, DoV:02]

- divide 2-D spectrum into slices with iterated tree-structured f-banks



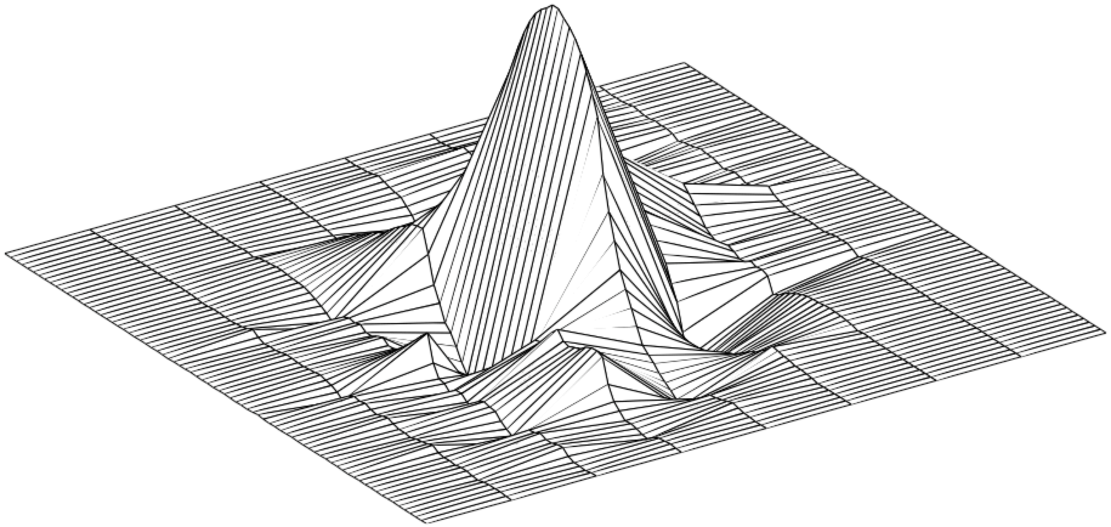
fan filters

quincunx sampling

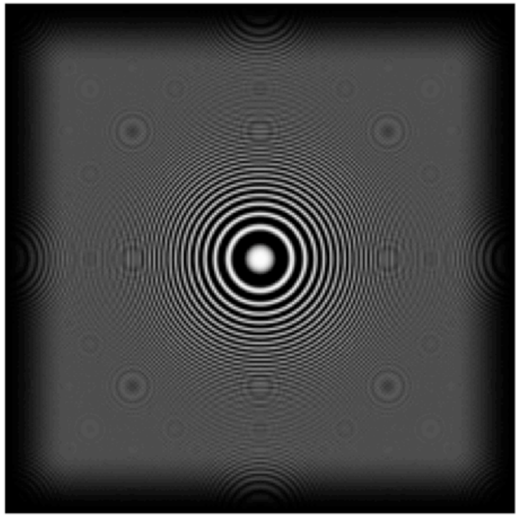


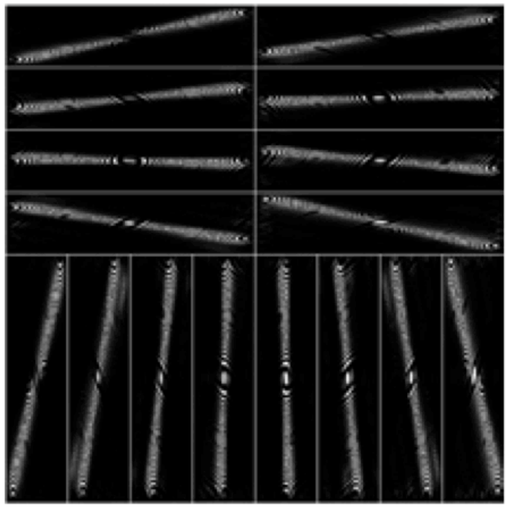
shearing

Iterated directional filter banks: efficient directional analysis



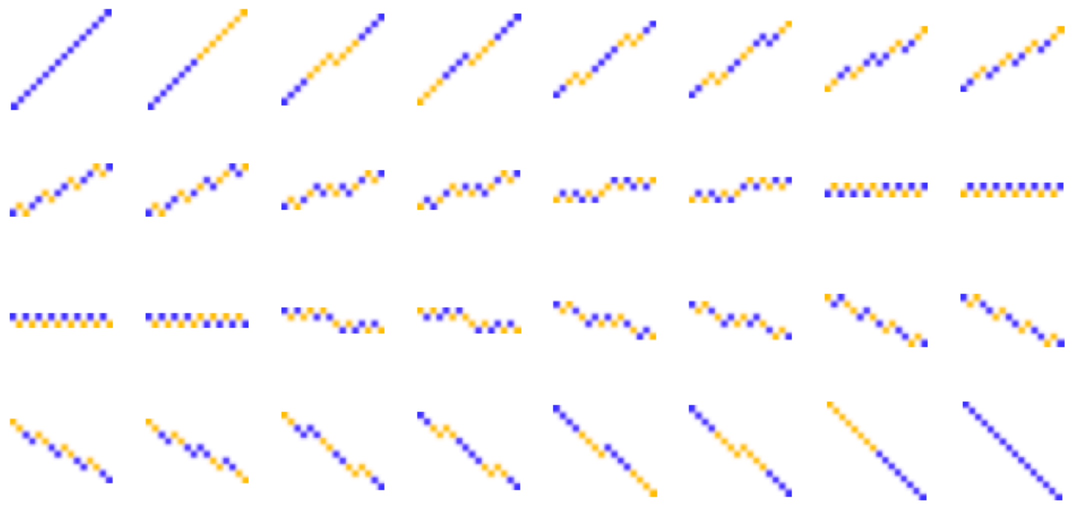
Example:





Example of basis functions

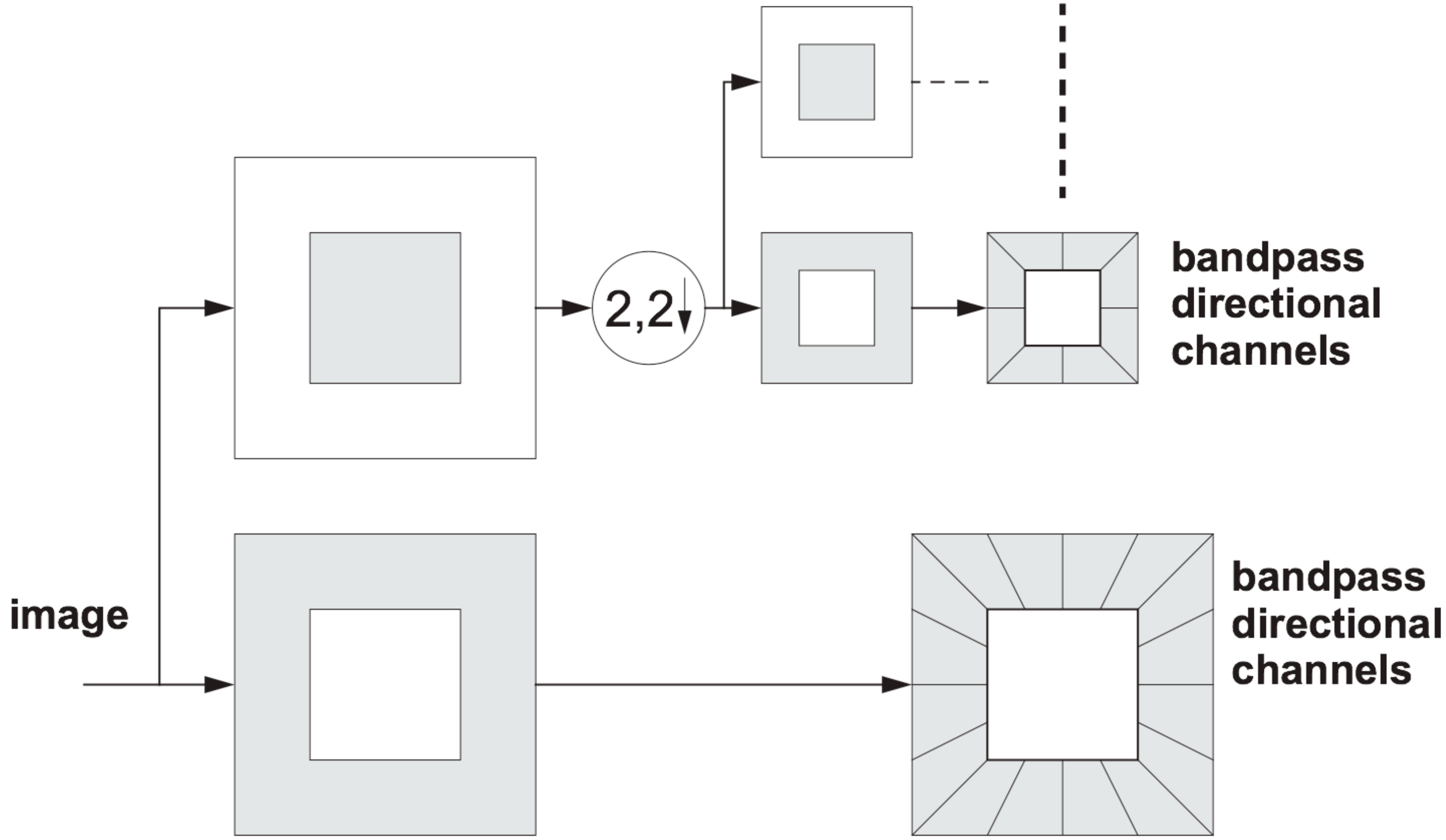
- 6 levels of iteration, or 64 channels
- elementary filters are Haar filters
- orthonormal directional basis
- 64 equivalent filters, below the 32 “mostly horizontal” ones are shown



This ressembles a “local Radon transform”, or radonlets!

- changes of sign (for orthonormality)
- approximate lines (discretizations)

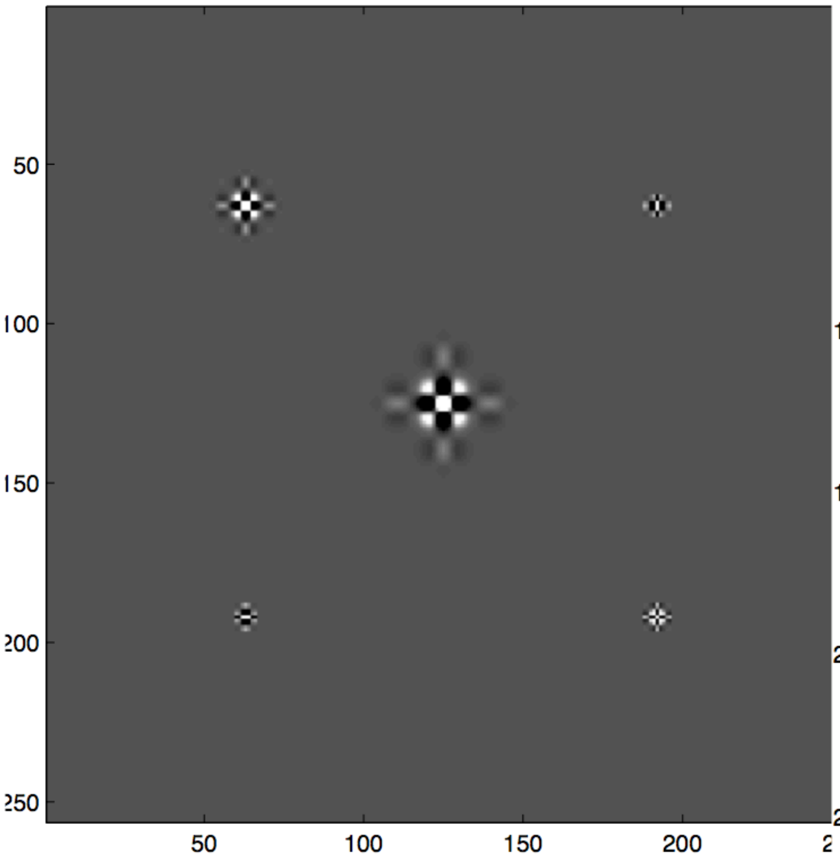
Adding multiresolution: use a pyramid!



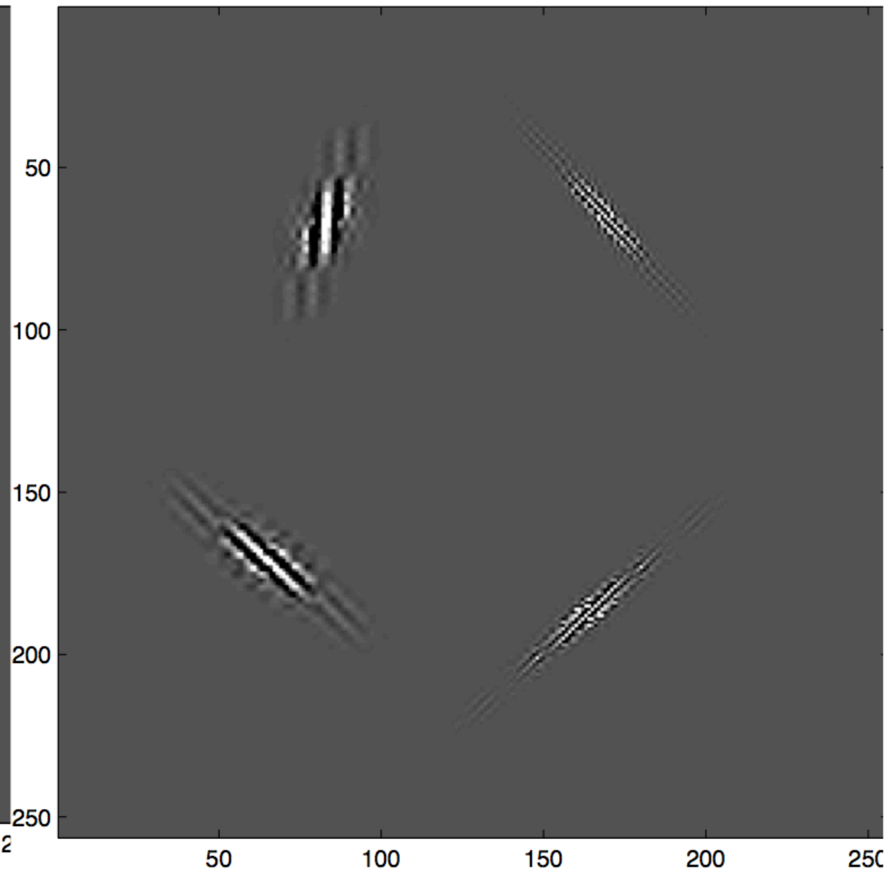
Result:

- “tight” pyramid and orthogonal directional channels => tight frame
- low redundancy < 4/3, computationally efficient

Basis functions: Wavelets versus contourlets



Wavelets

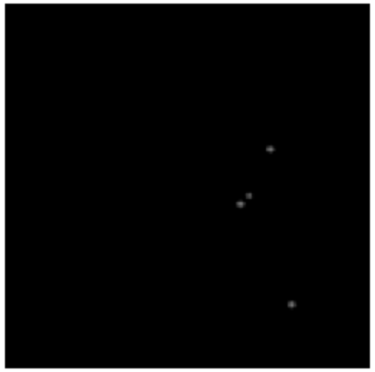


Contourlets

Approximation properties

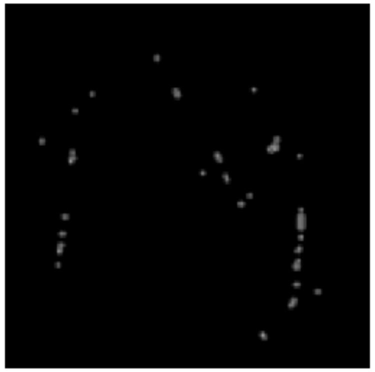
Wavelets

M = 4, MSE = 1.68e–4

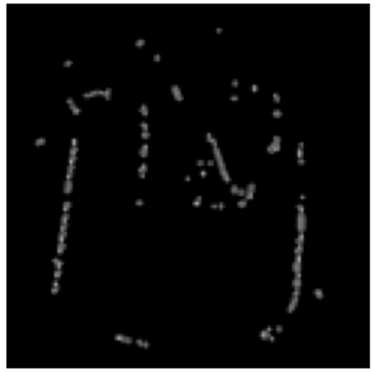


M = 16, MSE = 1.66e–4

M = 64, MSE = 1.60e–4

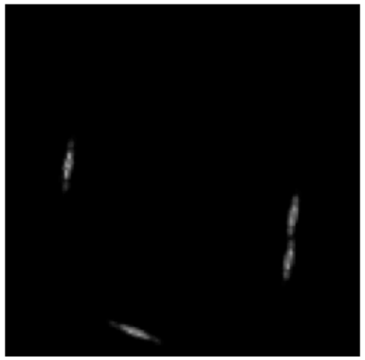


M = 256, MSE = 1.44e–4



Contourlets

M = 4, MSE = 1.68e–4



M = 16, MSE = 1.63e–4

M = 64, MSE = 1.55e–4

M = 256, MSE = 1.43e–4

An approximation theorem

Curvelets lead to optimal approximation, what about contourlets?

Result [M.Do:03]

Simple B/W image model with c^2 boundary

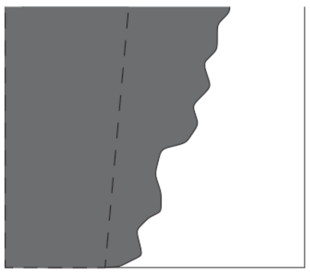
Contourlet with scaling $w \sim l^2$ and 1 directional vanishing moment

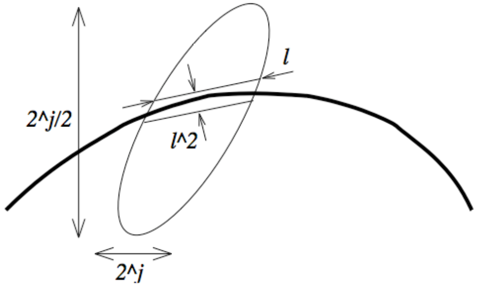
Then the M-term NLA satisfies

$$|f - f_{contourlet}|^2 \sim \frac{1}{M^2}$$

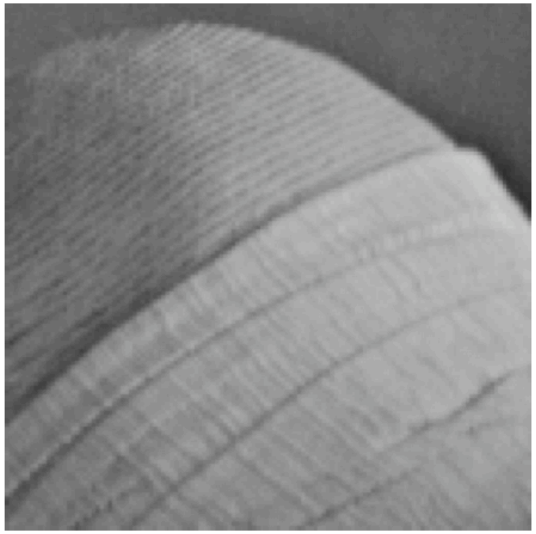
Proof (very sketchy...):

- Amplitude of contourlets ~ $2^{-3j/4}$ and coeffs ~$2^{-3j/4}\, l^3_{jkn}$
- Three types of coefficients (significant which match direction insignificant that overlap, and zero)
- levels 3J and J, respectively, leading to M ~ $2^{3J/2}$
- squared error can be shown to be ~ 2^{-3J}

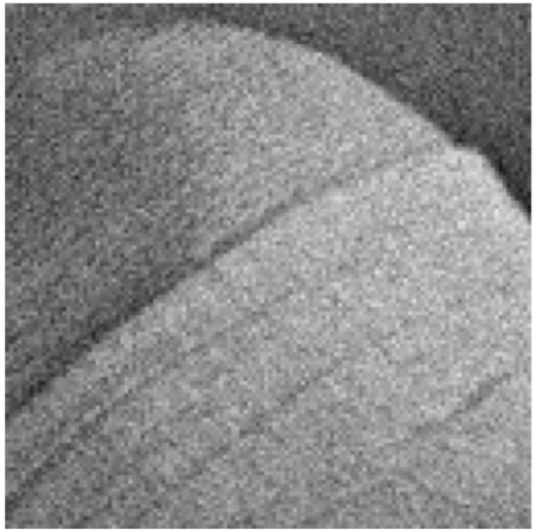




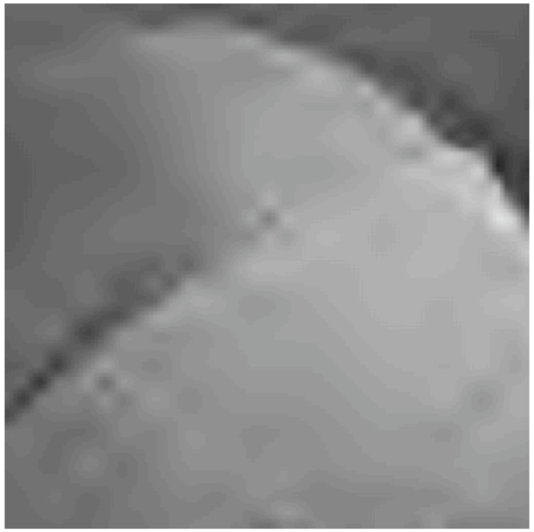
Example: denoising with contourlets



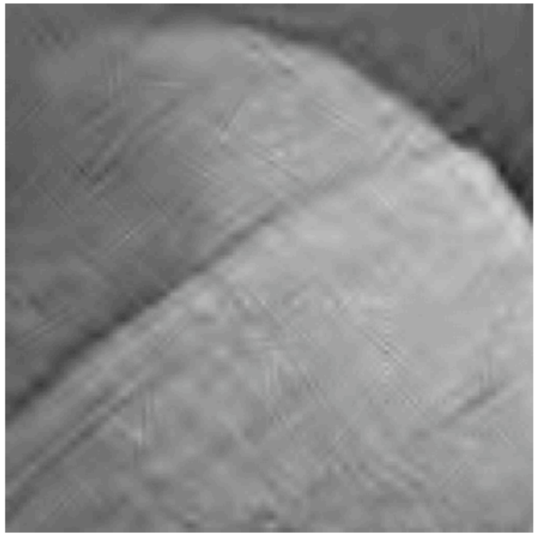
original



noisy



wavelet
13.8 dB



countourlets
15.4 dB

Outline

1. Introduction through History
2. Fourier and Wavelet Representations
3. Wavelets and Approximation Theory
4. Wavelets and Compression
5. Going to Two Dimensions: Non-Separable Constructions
6. Beyond Shift Invariant Subspaces: Finite Rate of Innovation
    - Shift-Invariance and Multiresolution Analysis
    - A Variation on a Theme by Shannon
    - A Representation Theorem
7. Conclusions and Outlook

Shift-Invariance and Multiresolution Analysis

Most sampling results require shift-invariant subspaces

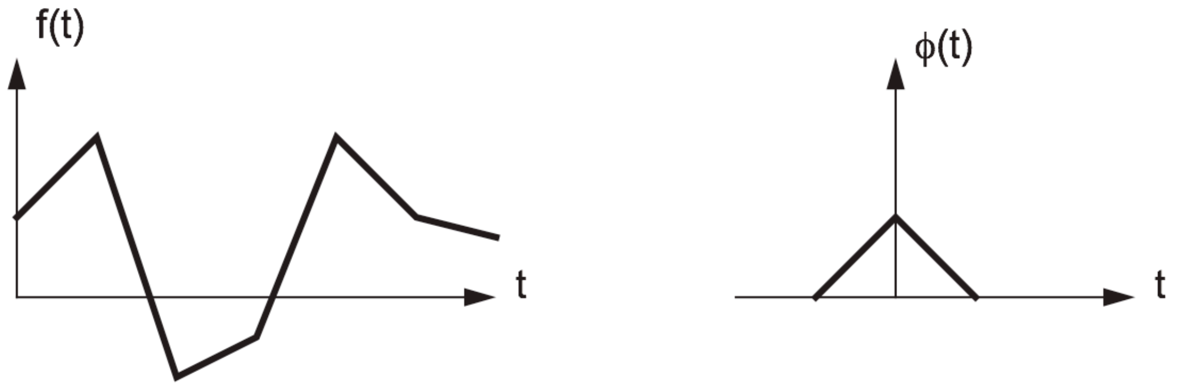
$$f(t) \in V \Leftrightarrow f(t - nT) \in V \qquad n \in Z$$

Wavelet constructions rely in addition on scale-invariance

$$f(t) \in V_0 \Leftrightarrow f(2^m t) \in V_{-m} \qquad m \in Z$$

Multiresolution analysis (Mallat, Meyer) gives a powerful framework.
Yet it requires a subspace structure.

Example: uniform or B-splines



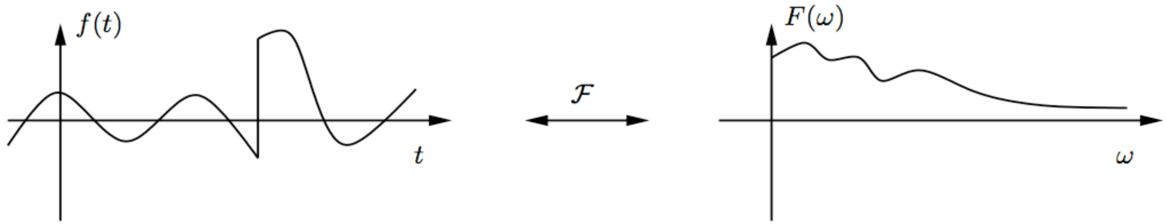
Question: can sampling be generalized beyond subspaces?

Note: Shannon BW sufficient, not necessary

A Variation on a Theme by Shannon

Shannon, BL case: $f(t) = \sum_{n \in Z} f(nT) sinc(t/T - n)$ or 1/T degrees of freedom per unit of time

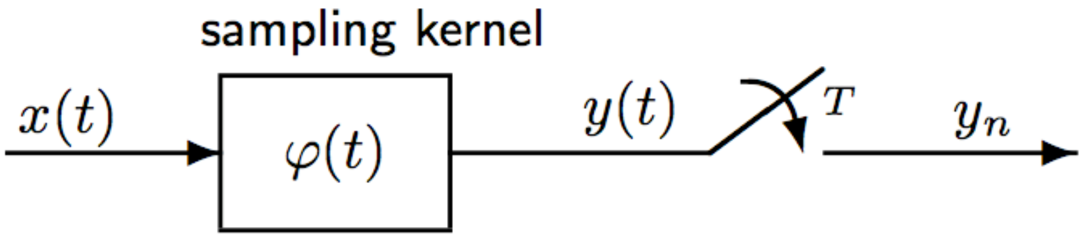
But: a single discontinuity, and no more sampling theorem...



Are there other signals with finite number of degrees of freedom per unit of time that allow exact sampling results?

=> Finite rate of innovation

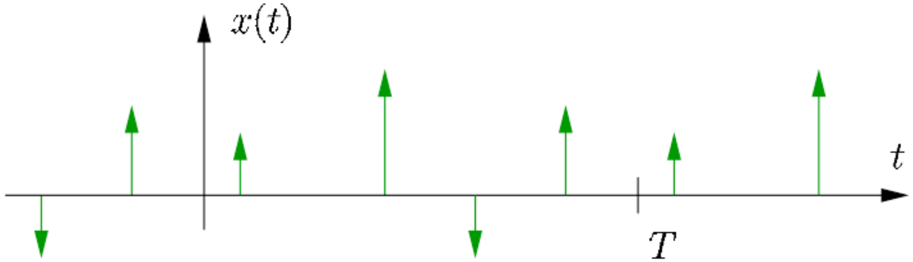
Usual setup:



x(t): signal, φ(t): sampling kernel, y(t): filtering of x(t) and y_n: samples

A Toy Example

K Diracs on the interval: 2K degrees of freedom. Periodic case:



$$x(t) = \sum_{n \in Z} \sum_{k=0}^{K-1} c_k \delta(t - t_k - n\tau) = \sum_{k=0}^{K-1} c_k \frac{1}{\tau} \sum_{m \in Z} e^{\frac{j2\pi m(t-t_k)}{\tau}}$$

Key: The Fourier series is a weighted sum of K exponentials

$$X[m] = \frac{1}{\tau} \sum_{k=0}^{K-1} c_k e^{\frac{-j2\pi m t_k}{\tau}}$$

Result: Taking 2k+1 samples from a lowpass version of BW-(2K+1) allows to perfectly recover x(t)

Method: Yule-Walker system, annihilating filter, Vandermonde system

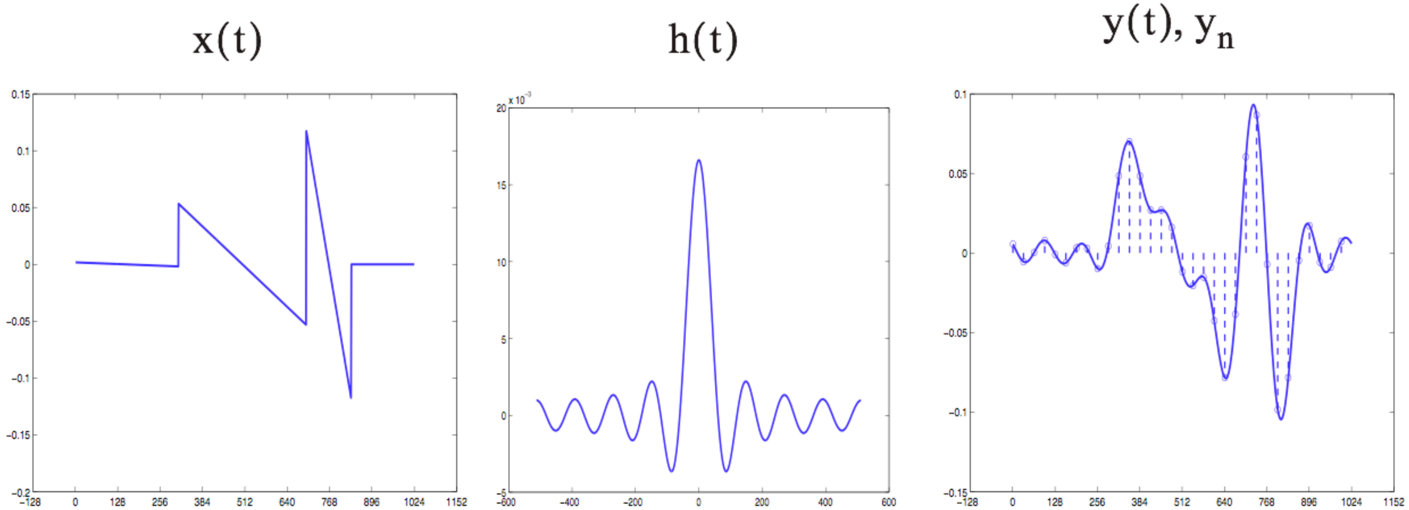
Note: Relation to spectral estimation and ECC (Berlekamp-Massey)

A Representation Theorem [VMB:02]

For the class of periodic FRI signals which includes

- sequences of Diracs
- non-uniform or free knot splines
- piecewise polynomials

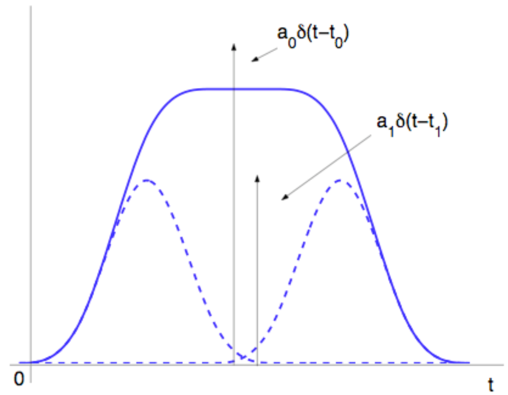
there exist sampling schemes with a sampling rate of the order of the rate of innovation which allow perfect reconstruction at polynomial complexity



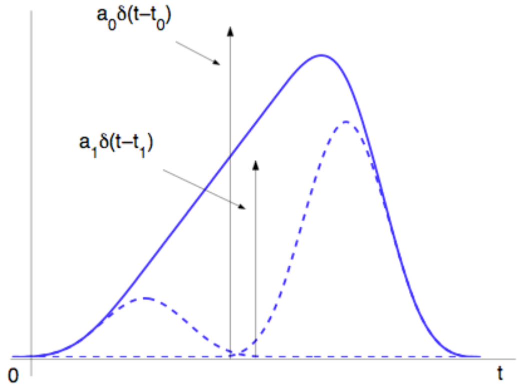
Variations: finite length, 2D, local kernels etc

A local algorithm for FRI sampling

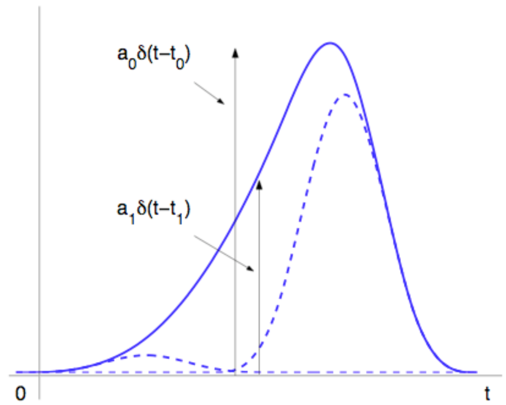
The return of Strang-Fix!



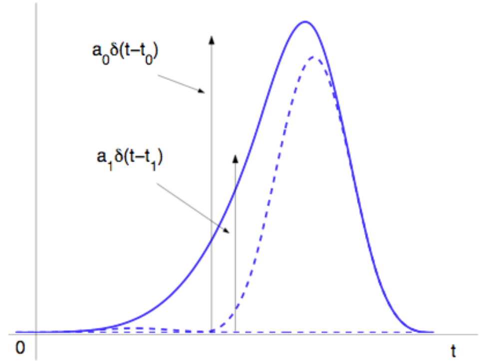
$$\sum_n y_n = a_0 + a_1$$



$$\sum_n n y_n = a_0 t_0 + a_1 t_1$$



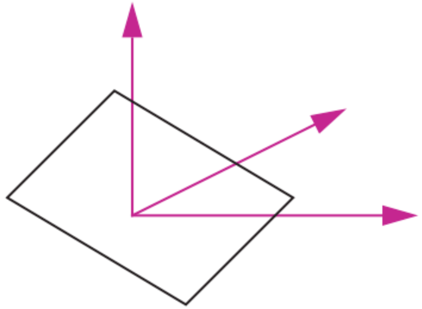
$$\sum_n n^2 y_n = a_0 t_0^2 + a_1 t_1^2$$

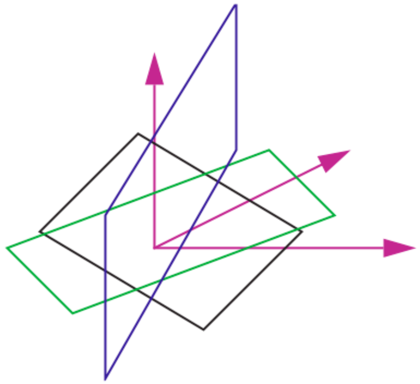


$$\sum_n n^3 y_n = a_0 t_0^3 + a_1 t_1^3$$

local, polynomial complexity reconstruction, for diracs and piecewise polynomials

Conclusions





Wavelets and the French revolution

- too early to say?
- from smooth to piecewise smooth functions

Sparsity and the Art of Motorcycle Maintenance

- sparsity as a key feature with many applications
- denoising, inverse problems, compression

LA versus NLA:

- approximation rates can be vastly different!

To first order, operational, high rate, D(R)

- improvements still possible
- low rate analysis difficult

Two-dimensions:

- really harder! and none used in JPEG2000...
- approximation starts to be understood, compression mostly open

Beyond subspaces:

- FRI results on sampling, many open questions!

Outlook

Do we understand image representation/compression better?

- high rate, high resolution: there is promise
- low rate: room at the bottom?

New images

- plenoptic functions (set of all possible images)

- non BL images (FRI?)
- manifolds, structure of natural images

Distributed images

- interactive approximation/compression
- SW, WZ, DKLT...

Why Image Representation Remains a Fascinating Topic…

A lone student standing
in front of four tanks.

Publications

For overviews:

- D.Donoho, M.Vetterli, R.DeVore and I.Daubechies, Data Compression and Harmonic Analysis, IEEE Tr. on IT, Oct.1998.
- M. Vetterli, Wavelets, Approximation and Compression, IEEE Signal Processing Magazine, Sept. 2001.

Research papers:

- See http://lcavwww.epfl.ch/

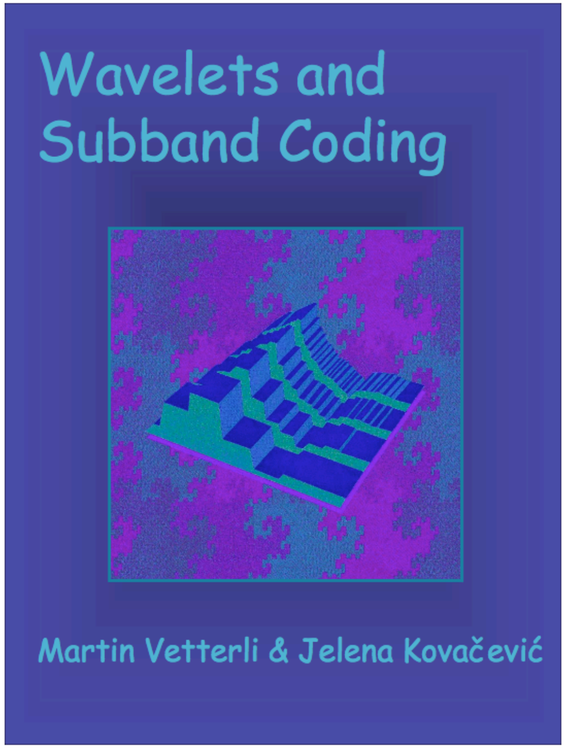
Recent Theses (online):

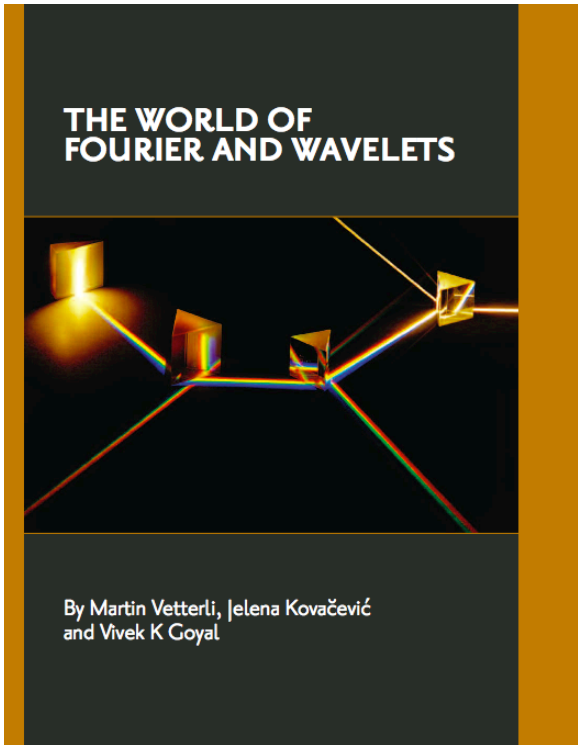
- C.Weidmann on rate-distortion analysis of NLA
- P.L.Dragotti on wavelet footprints
- Minh Do on contourlets
- P.Marziliano on new sampling theorems of non BL fcts

Publications

Books:

- M.Vetterli and J.Kovacevic, Wavelets and Subband Coding, Prentice-Hall (1995), Open Access (2007).
- M.Vetterli, J.Kovacevic and V.Goyal, The World of Fourier and Wavelets: Theory, Algorithms and Applications (2008)





Thank you for your attention!