On Wavelet-Based Image Compression and Beyond

Pier Luigi Dragotti

Communications and Signal Processing Group

Imperial College London

Acknowledgements

- Prof. Martin Vetterli (Swiss Federal Institute of Technology & UC Berkeley)
- Prof. Minh Do (University of Illinois at Urbana-Champaign)
- Nicolas Gehrig (Imperial College London)
- Rahul Shukla, Vladan Velisavlievic and Baltasar Beserull-Lozano (EPFL)

Outline

- Motivation: Importance and limit of the wavelet transform
- The 1-D case
  - Signals of interests and wavelet representations
  - Wavelet Footprints
  - Approximation and Compression
- From 1-D to 2-D
  - Quad-Tree Decomposition
  - Directional Wavelets
- Distributed Source Coding and Sensor Networks

A Success Story

Wavelets are in the new image compression standard (JPEG2000)

Original Lena Image (256×256 pixels)

JPEG (Compression Ratio 43:1)

JPEG2000 (Compression Ratio 43:1)

Note: images courtesy of dspworx.com

Motivation

Sparse Representations

Wavelets are successful because they provide sparse representations of images.

Cameraman

Wavelet decomposition

8% of coeff. (SNR=25.3dB)

Motivation

Wavelet methods (simple and efficient):
Images are decomposed in the wavelet basis and larger coefficients are kept.
Example: denoising

SNR=17dB

SNR=20.2dB

Wavelets are also useful for interpolation, classification...

Motivation

But *a* nobel prize told us that

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

Lena Image

JPEG2000 (35Kbits)

Stare at the original image for 10 seconds. You have just taken in 200bits!

Motivation

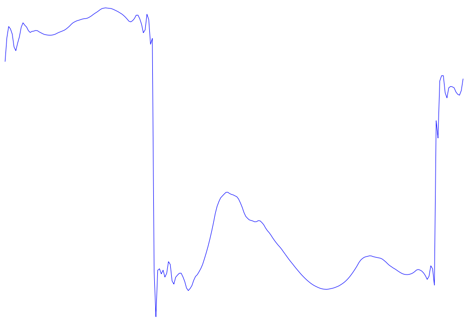
Limitations of the wavelet transform:

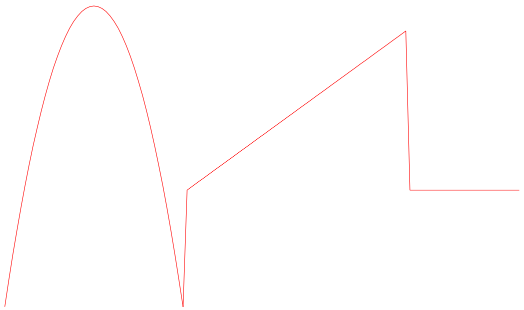
- do not take advantage of dependency across scales,
- do not take advantage of geometric regularity.

The 1-D case. Signals of interest

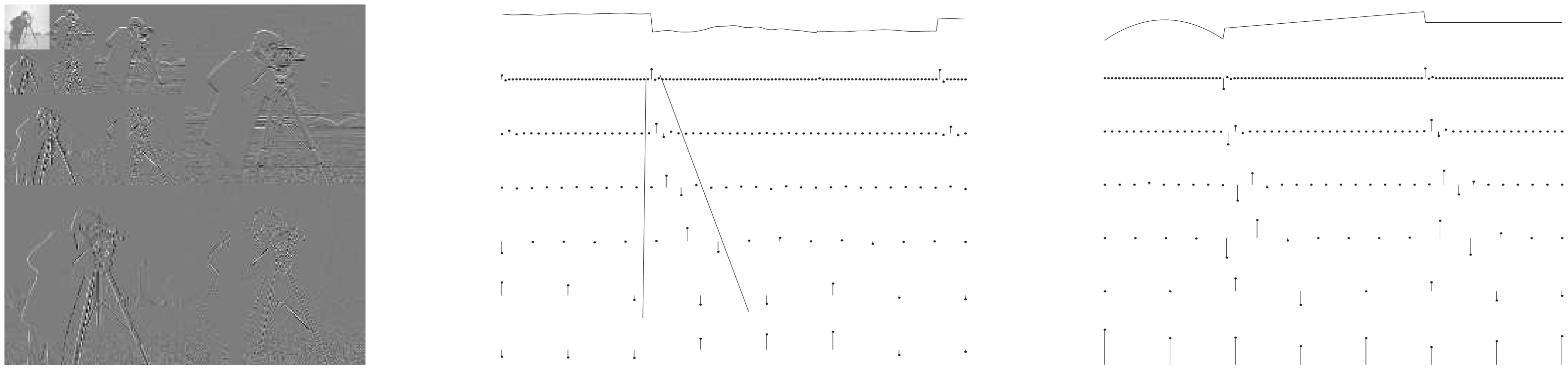
Signals of interest:

- piecewise smooth signals (that is, signals made of Lipshitz-α pieces),
- piecewise polynomial signals.





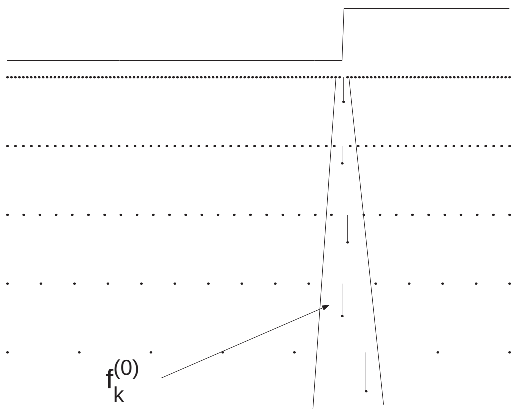
Signals of Interest and Wavelet Representations



- Wavelet coefficients around smooth parts of the signal are small and have fast decay ($\sim 2^{-j(\alpha+1/2)}$).
- Wavelet coefficients around polynomial parts of the signal are exactly zero.
- Discontinuities generate a finite number of large wavelet coefficients.

Piecewise constant discontinuity

Insight: all the wavelet coefficients across scales generated by a step discontinuity are dependent. They have only one degree of freedom.

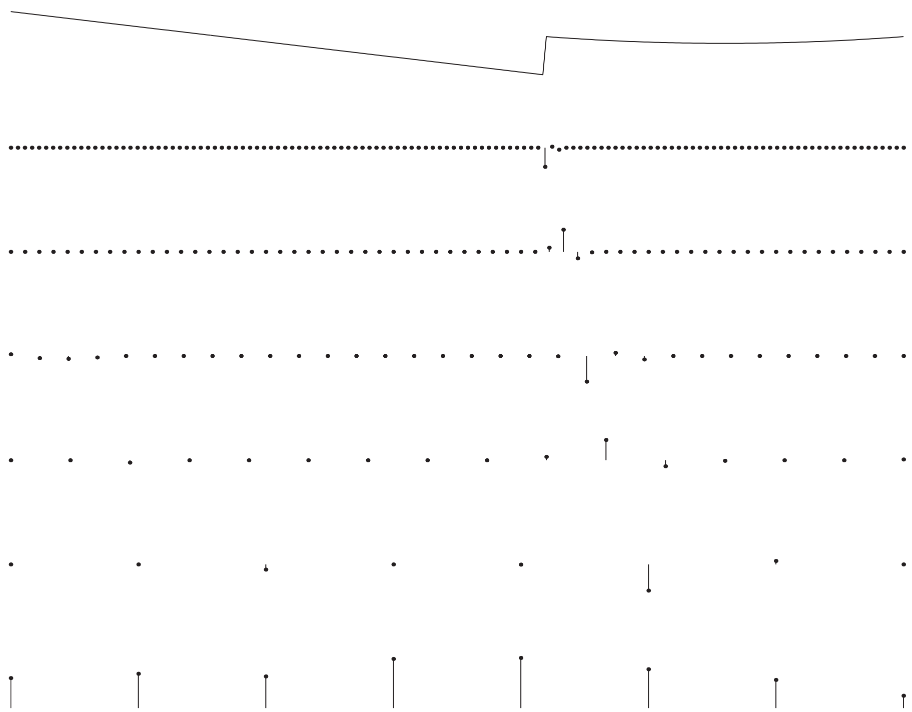


Footprint: Call footprint $f_k^{(0)}$ the norm 1 scale-space vector obtained by gathering together all the wavelet coefficients in the cone of influence of k and then imposing $\| f_k^{(0)} \| = 1$.

The wavelet coefficients Y generated by any step discontinuity at k are given by:

$$Y = \alpha^{(0)} f_k^{(0)}, \quad \text{where} \quad \alpha^{(0)} = \langle Y, f_k^{(0)} \rangle$$

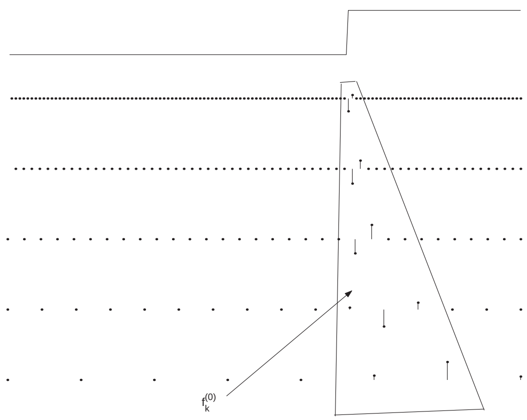
Piecewise polynomial discontinuity

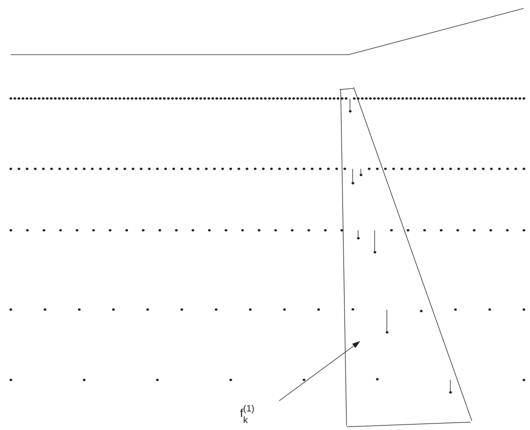


The set of wavelet coefficients generated by a discontinuity are dependent: they lie on a subspace of dimension $D+1$ (D maximum degree of any polynomial in the signal).

Piecewise polynomial discontinuity

Create $D + 1$ footprints $f_k^{(d)}$ related to a discontinuity of order $0, 1, .., D$, such that: $\|f_k^{(d)}\| = 1$ and $< f_k^{(i)}, f_k^{(j)} >= \delta_{ij}$.





The wavelet coefficients generated by any piecewise polynomial discontinuity at k are given by:

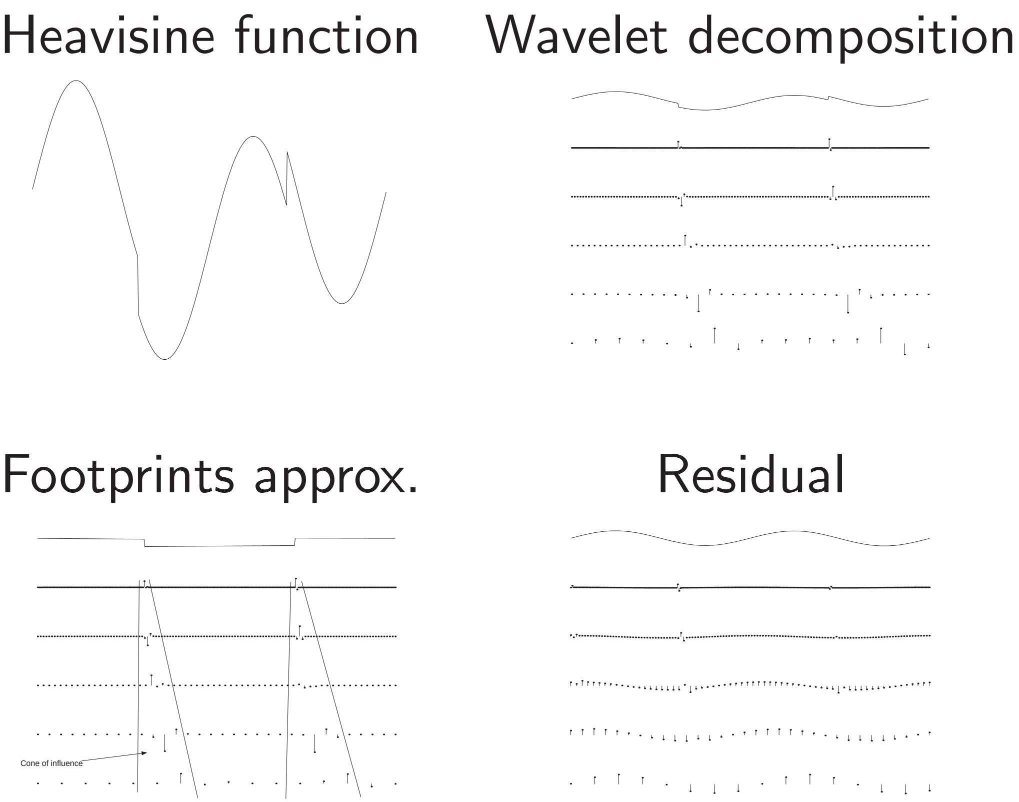
$$Y = \sum_{d=0}^{D} \alpha^{(d)} f_k^{(d)},$$

$$\alpha^{(d)} = \langle Y, f_k^{(d)} \rangle.$$

Footprints Dictionary

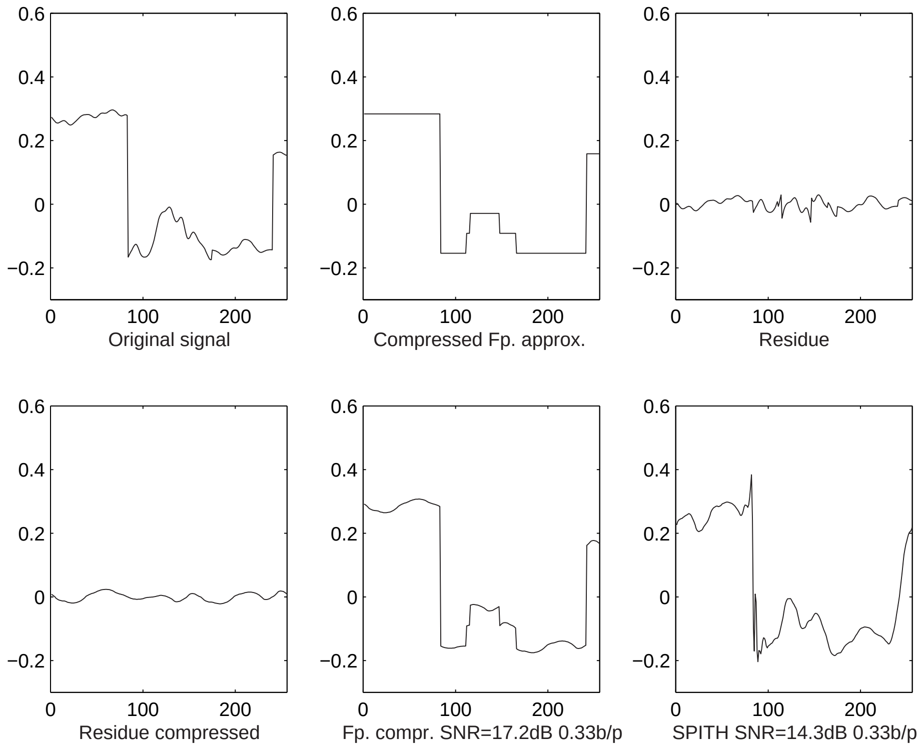
- Create $D + 1$ different footprints for each discontinuity location $k \in [0, N - 1]$.
- Call $\mathcal{D} = \{f_k^{(d)}, k \in [0, N - 1], d = 0, 1.., D\}$ the dictionary containing all footprints.
- The dictionary $\mathcal{D}$ can provide a compact representation of any piecewise polynomial signal.
- The dictionary $\mathcal{D}$ is unconditional for the class of piecewise polynomial signals (Dragotti-Vetterli:03).

Footprints in Action



Theorem (Dragotti-Vetterli:03): Given is a piecewise smooth signal $f(t)$ with α-Lipschitz regular pieces. There exists a piecewise polynomial signal $p(t)$ with pieces of maximum degree $p = \lfloor \alpha \rfloor$ such that the difference signal $g(t) = f(t) - p(t)$ is uniformly Lipschitz α over $[0, T]$.

Applications: Compression



Compression with Footprints: SNR=17.2dB, 0.33b/p.

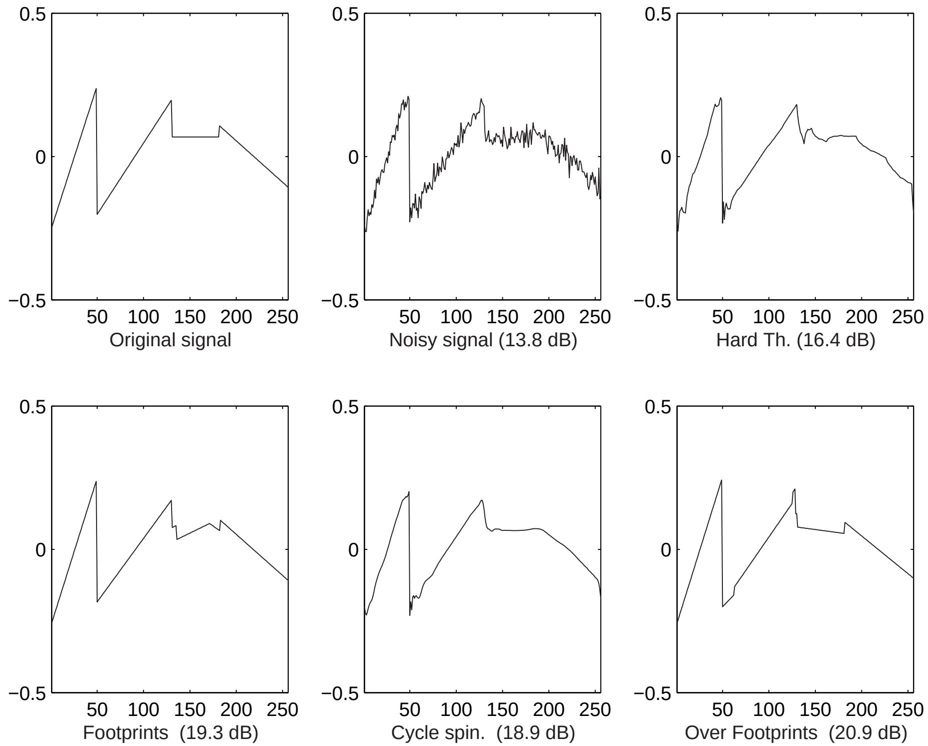
Compression with SPIHT: SNR=14.3dB, 0.33b/p.

Approximation and Compression

More formally:

- Non-linear approximation of piecewise smooth signals
  - With Fourier $MSE \sim M^{-1}$
  - With Wavelets $MSE \sim M^{-2\alpha}$
- Compression of piecewise smooth signals
  - With Fourier $D(R) \leq c_0 R^{-1}$
  - With Wavelets $D(R) \leq c_1 R^{-2\alpha} + c_2 \sqrt{R} 2^{-c_2 \sqrt{R}}$ ([CohenDGO:02])
  - With Footprints $D(R) \leq c_3 R^{-2\alpha} + c_4 2^{-c_5 R}$ ([DragottiV:03])
- Compression of piecewise polynomial signals
  - With Wavelets $D(R) \leq c_2 \sqrt{R} 2^{-c_2 \sqrt{R}}$
  - With Footprints $D(R) \leq c_4 2^{-c_5 R}$

Applications: Denoising



Cycle-spinning (Coifman-Donoho 1995): SNR=18.9dB
Cycle-spinning with footprints: SNR=20.9dB.

From 1-D to 2-D

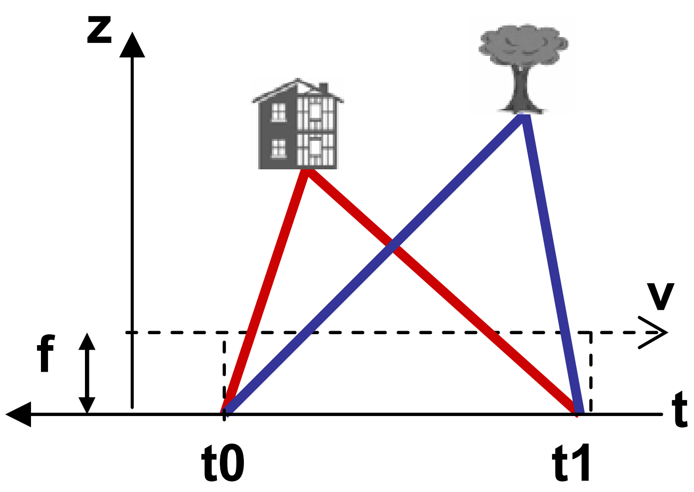
With footprints we can achieve optimal performance. But how many visual signals are 1-D?

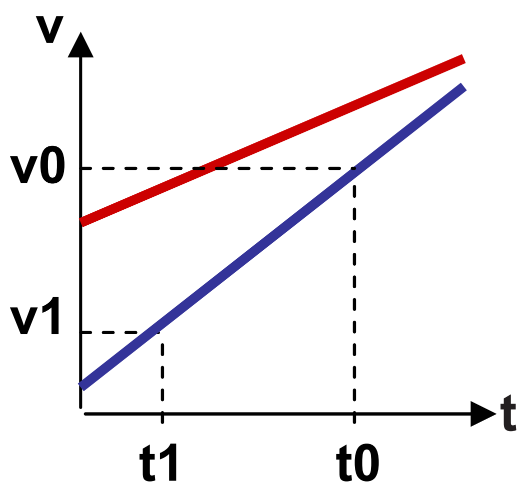
- Grayscale images 2-D
- Color images 3-D
- Video sequences 3-D+1=4-D
- Plenoptic Function 7-D...

Thus, this is not the end of the story.

Intermezzo: The Plenoptic Function

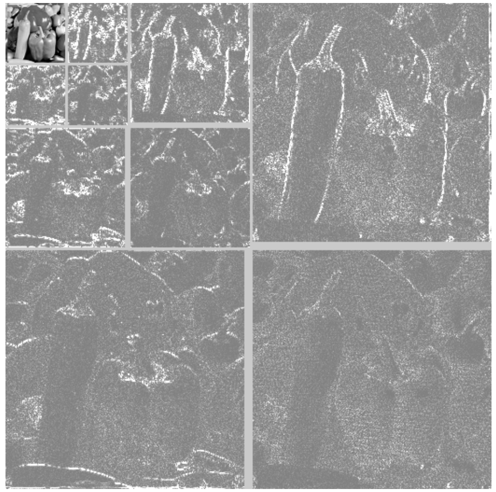
Adelson and Bergen change the rules of the game [AdelsonB:91]
'Assume that one is free to take photographs of a visual scene at any possible position, angle and time. Such a complete representation of that scene can be parameterized by a single function called the **Plenoptic Function**'





The images or the video sequences we play with are just particular realizations of the Plenoptic Function.

From 1-D to 2-D (continued)



Wavelets (separable transforms) have some limitations in higher dimensions.

- In images, the action is on the edges,
- Wavelet schemes fail to recognize that the boundary is smooth,
- New scheme requires geometrical processing.

People working on 'truly' 2-D representations

- Ridgelets and Curvelets (Candès, and Donoho)
- Contourlets (Do and Vetterli)
- Wedgelets and Beamlets (Donoho)
- Wedgeprints (Baraniuk)
- Edgeprints (Dragotti and Vetterli)
- Edge Adapted multiscale Transform (Cohen)
- Improved quad-tree decompositions (Shukla, Dragotti, Do and Vetterli)
- Bandelets (Le Pennec and Mallat)
- Complex Wavelets (Kingsbury)
- Discrete Directional Wavelets (Velisavlievic, Beferull-Lozano , Vetterli and Dragotti)

Image Model

- Two-dimensional piecewise smooth or polynomial functions with smooth or polynomial boundaries.

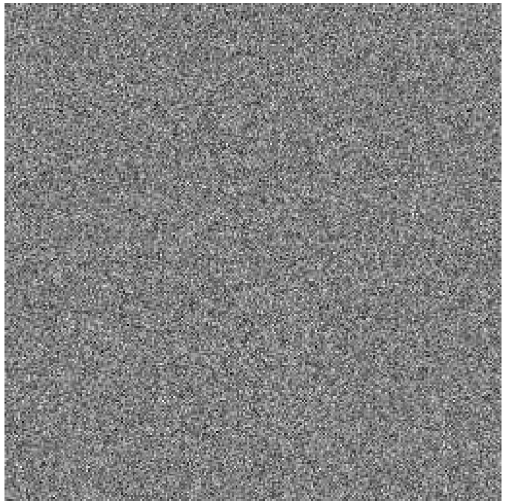
- "Oracle" performances

  - Boundary is a C^p curves: $D(R) \sim R^{-p}$
  - Boundary is a polynomial: $D(R) \sim 2^{-cR}$

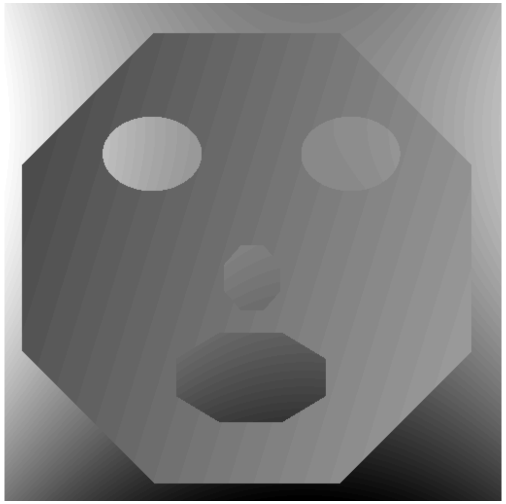
- Wavelet performance

  - **Regardless** of the smoothness of the boundary: $D(R) \sim \log(R)R^{-1}$

How far are we from Lena?



Gauss-Markov Model

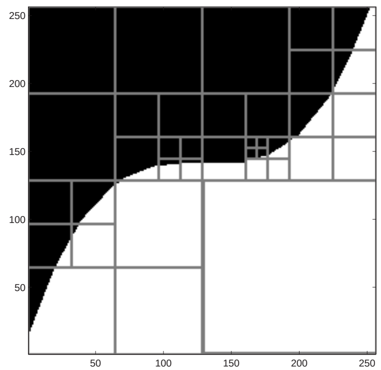


Piecewise Polynomial Model

Lena

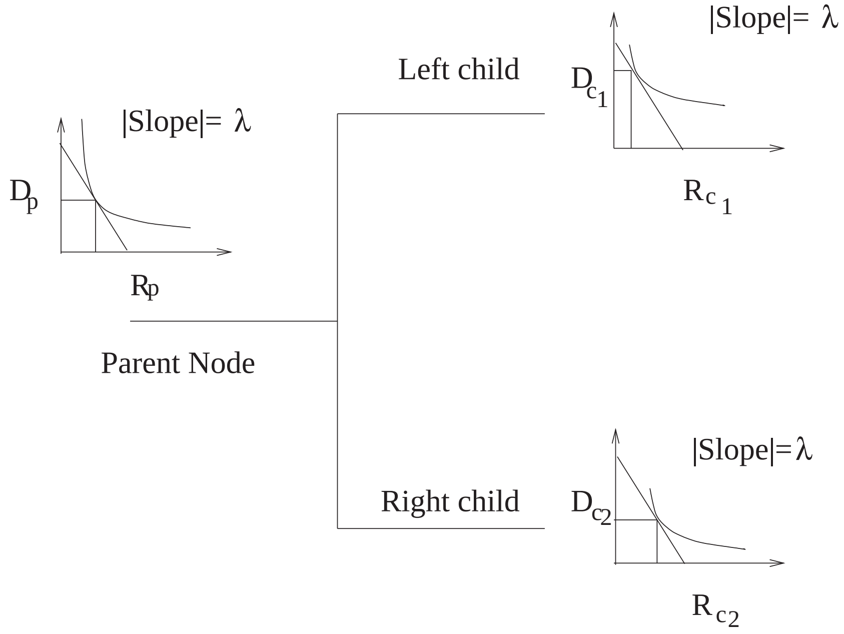
Quadtree Algorithms

- Prune Quadtree Algorithm.
  - Quadtree segment into dyadic squares.
  - Code each block with a "geometrical tile" made of two polynomial pieces divided by a linear discontinuity.
  - Prune the tree to minimize $D + \lambda R$.



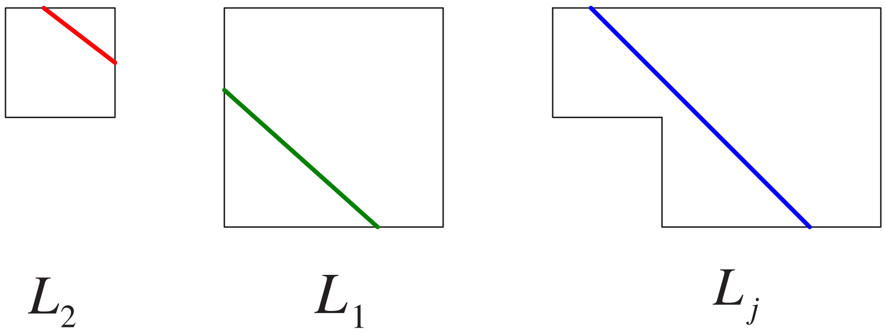
- Prune-Join Quadtree Algorithm with Joint Coding.
  - Find the optimal tree using the Pruned Quadtree Algorithm.
  - Code neighbor segments with "similar" parameters jointly.

To prune or not to prune



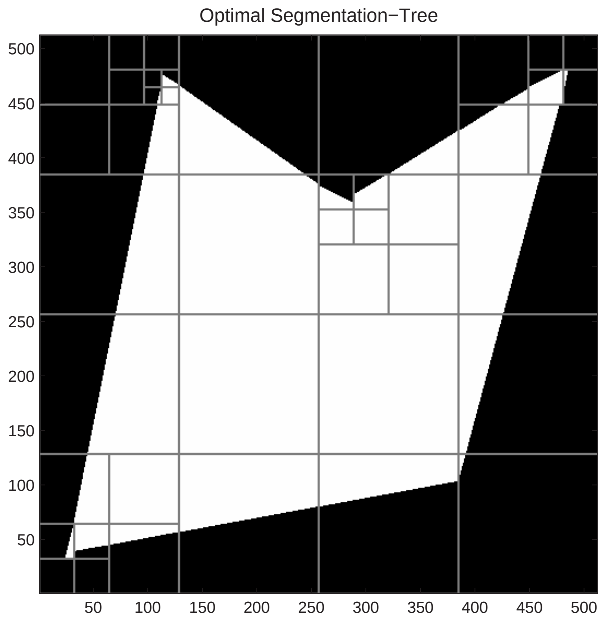
Prune if: $(D_{c_1} + D_{c_2}) + \lambda(R_{c_1} + R_{c_2}) > D_p + \lambda R_p$

To join or not to join

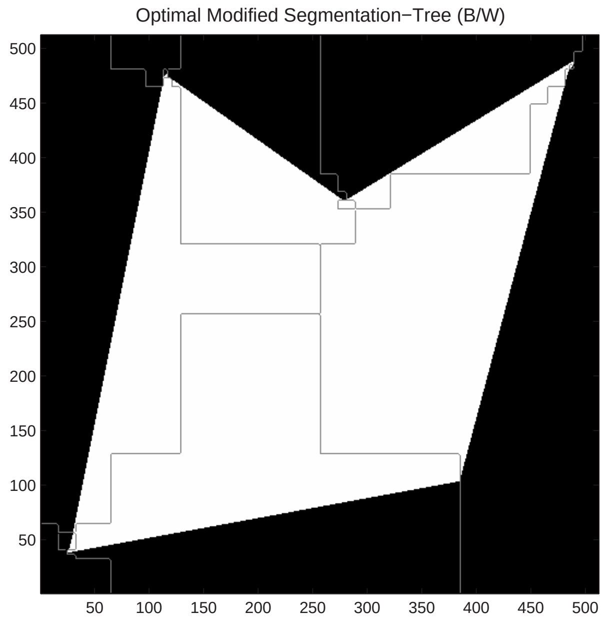


Join if: $L_J(\lambda) < L_1(\lambda) + L_2(\lambda)$ where $L(\lambda) = D + \lambda R$.

Example

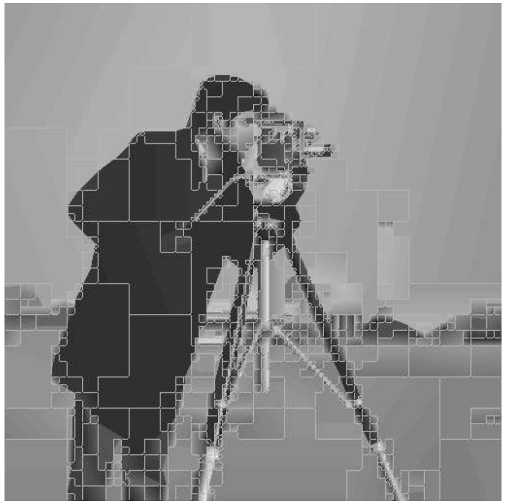


$$D(R) \sim 2^{-c\sqrt{R}}$$



$$D(R) \sim 2^{-cR}$$

Quad-Tree vs Jpeg 2000



Prune-Join Quad-Tree Decomposition

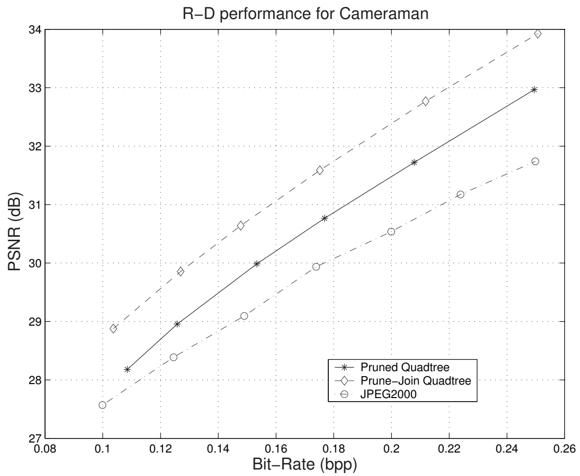
P-J Tree PSNR 28.9dB, 0.11bpp

JPEG2000 PSNR 27.8dB 0.11bpp

Quad-Tree vs Jpeg 2000

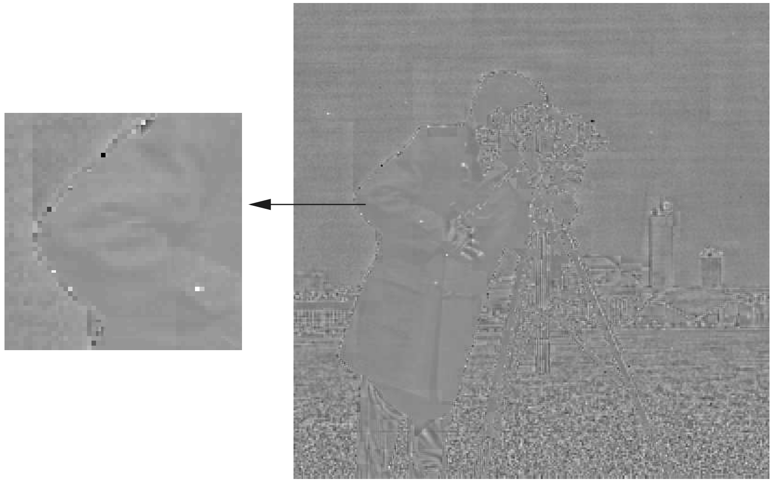
PSNR=30.68dB, 0.15bpp

PSNR=29.21dB, 0.15bpp

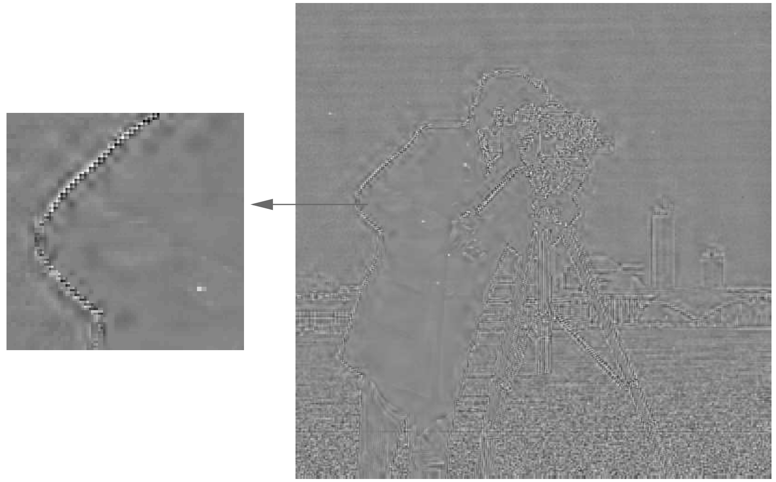


Quad-Tree vs JPEG2000

Quad-Tree vs Jpeg 2000



Prune-Join Quad-Tree

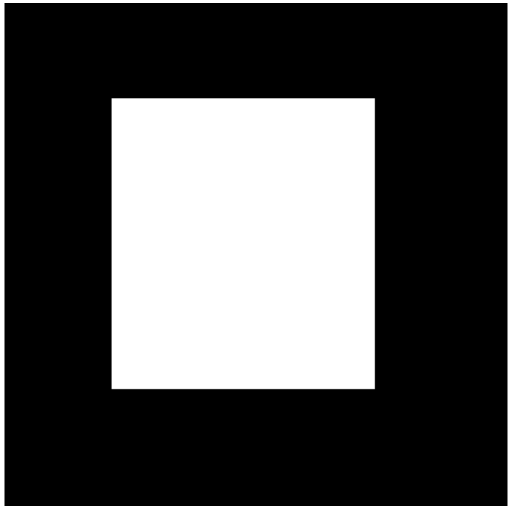


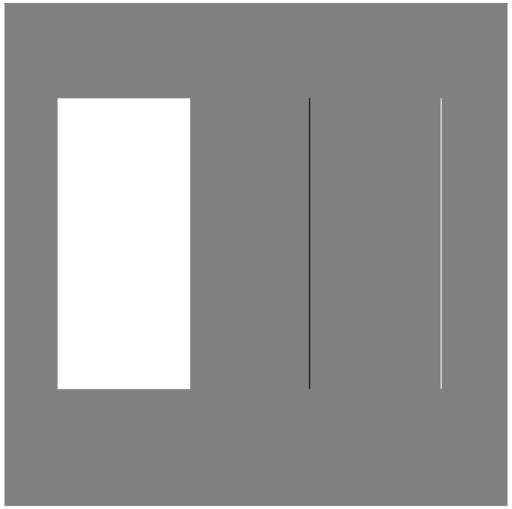
JPEG2000

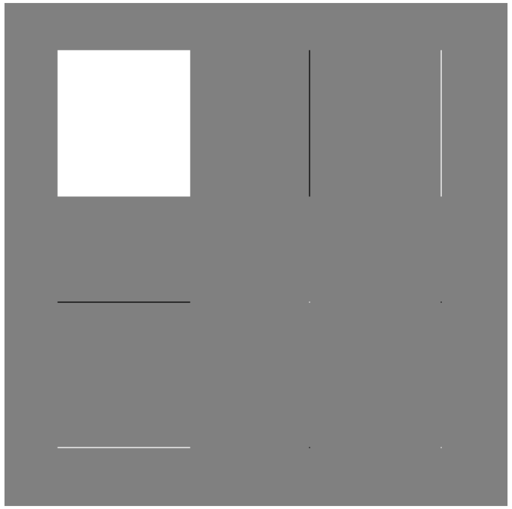
Directional Wavelets

Target:

- Keep the simplicity of 1-D processing,
- Focus on discrete signals to lead to algorithmic implementations.



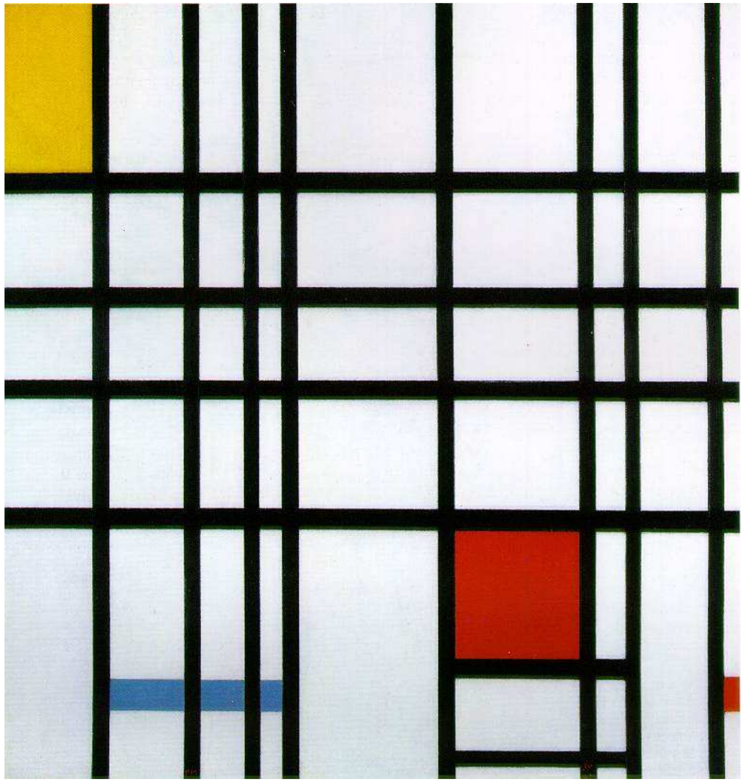




The separable wavelet transform is good at isolating horizontal and vertical edges.

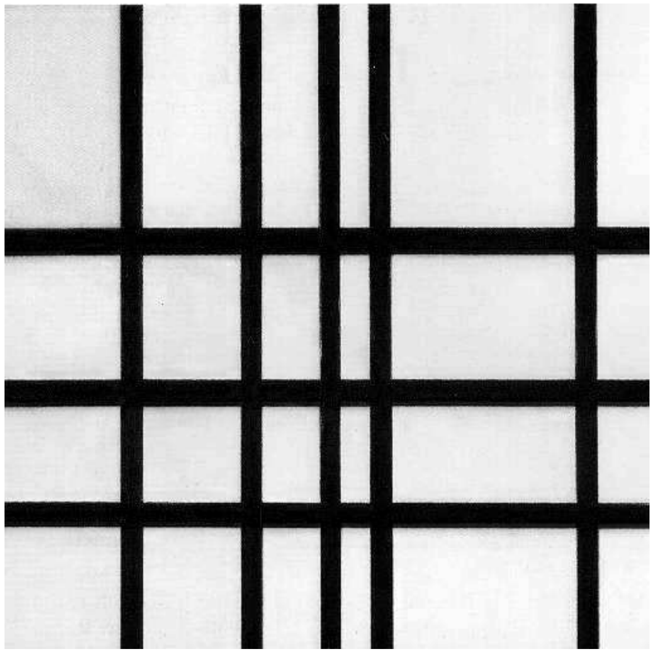
WebMuseum, Paris

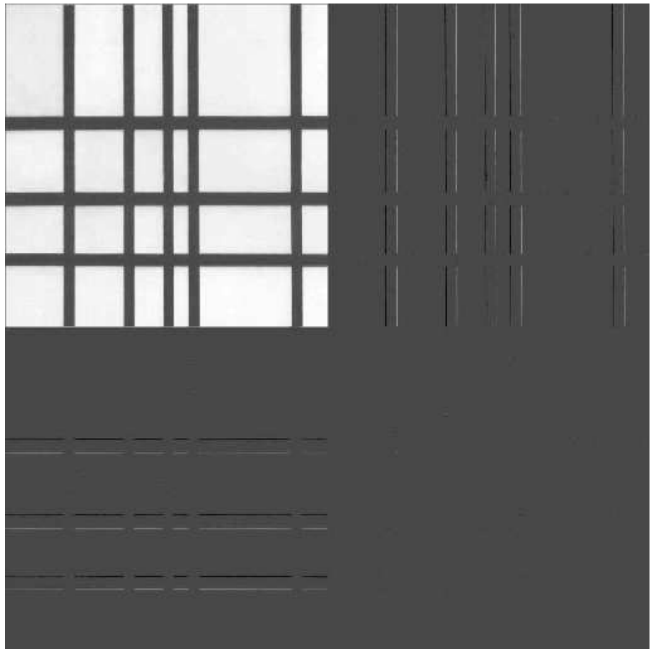
Mondrian, Piet



WebMuseum, Paris

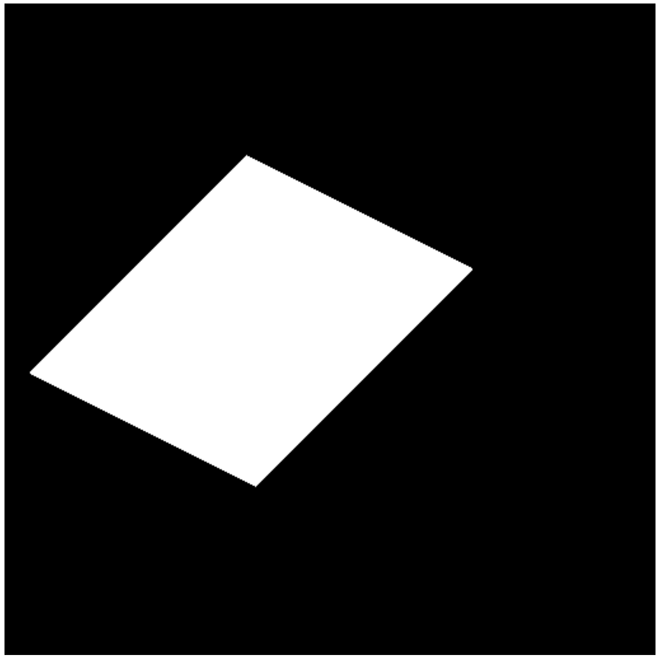
Wavelet decomposition of Mondrian masterpiece.

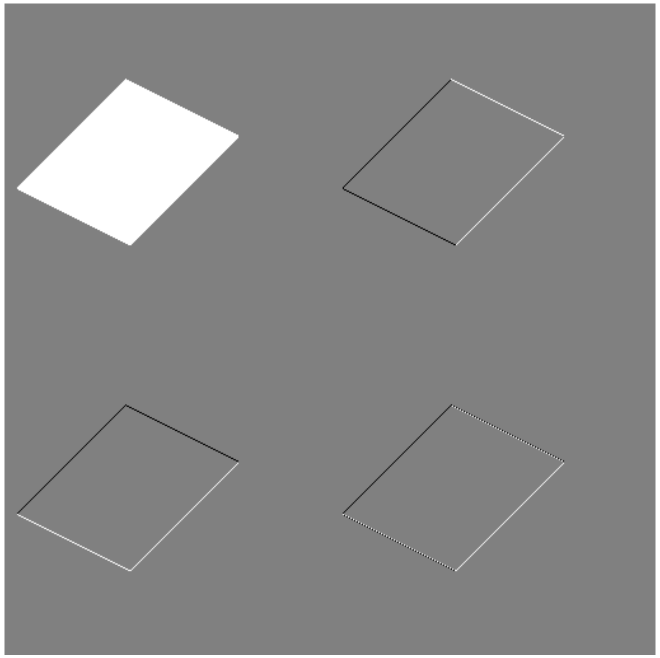




Directional Wavelets

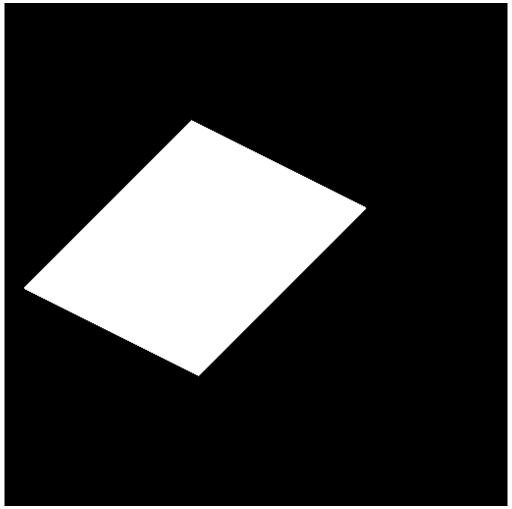
...but, it fails if directions are not horizontal or vertical.

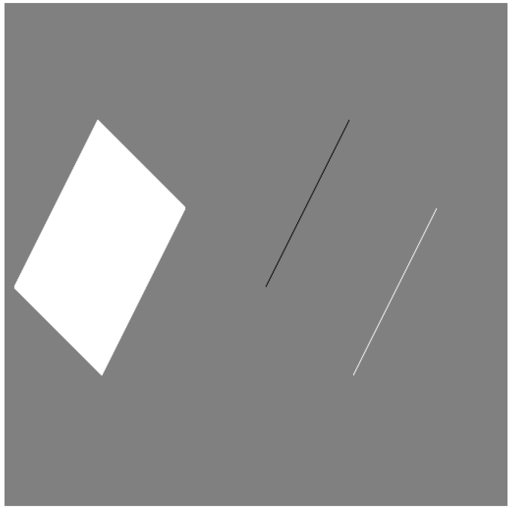


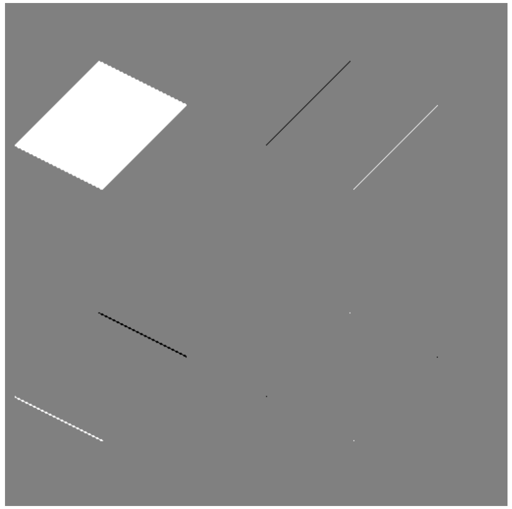


Directional Wavelets

Directional wavelets are as simple as traditional wavelets, but provide compact representations of edges along different directions.



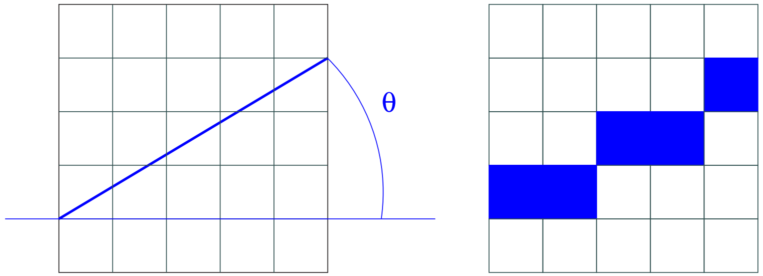




Directional Wavelets

Directional transforms. Basic elements.

- Definition of a digital line
$$y[n] = \lfloor kx[n] \rfloor + \lfloor B \rfloor$$
(Each pixel belongs to exactly one line for a chosen slope)



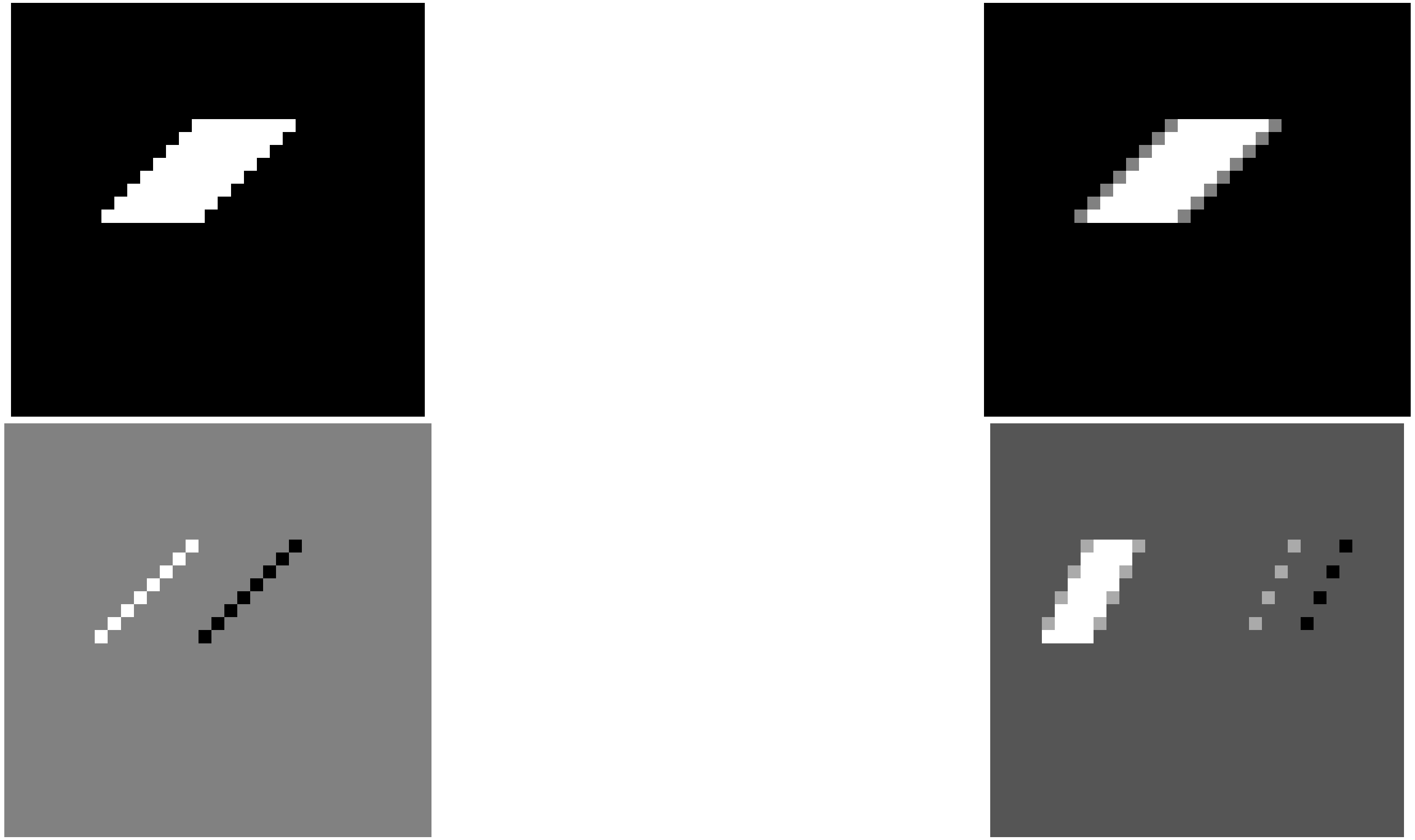
- Apply a decimated or undecimated wavelet transform along the digital line
- Iterate the process along either the same or a different direction.

This leads to a wide range of different multi-directional bases or frames

- Multi-directional bases useful for compression,
- Multi-directional frames useful for denoising.

Multi-directional Bases

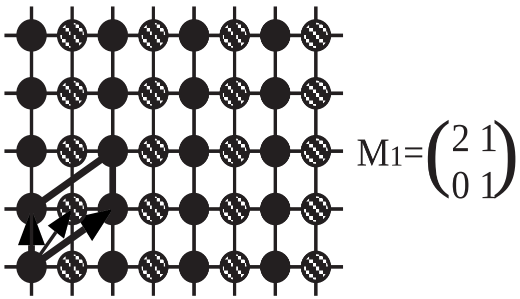
Subsampling is critical to obtain a nice basis. Recall that we want to annihilate lines $N \longrightarrow \log_2 N$.



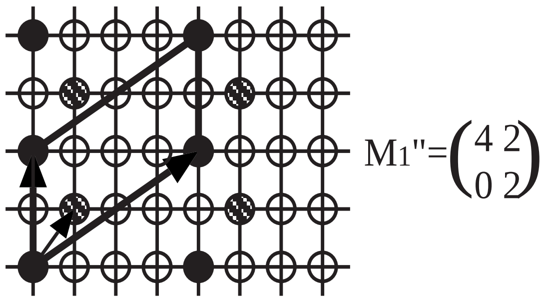
Multi-directional Bases

Resort to lattice theory.

- Choose two initial directions $r_1 = t_1/m_1, \qquad r_2 = t_2/m_2.$
- Construct the corresponding lattice $M_1 = \begin{pmatrix} m_1 & t_1 \\ m_2 & t_2 \end{pmatrix}$
- Number of cosets equal to $|\det(M_1)|$.
- Apply the 1-D transform along the two directions but threat each coset independently.
- Subsampling: $M_2 = 2M_1$.
- Iterate on the same directions or along different directions.



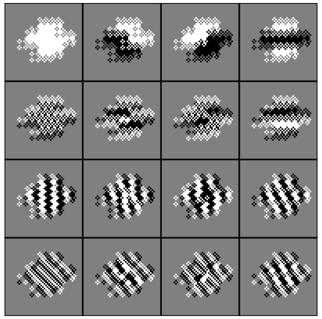
(a)



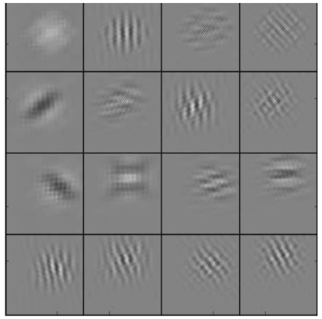
(b)

Multi-Directional Bases

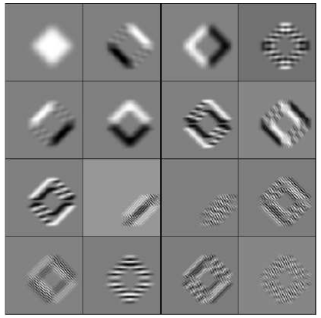
We still have some problems with the regularity.



Haar
orthogonal
easy to invert
non-smooth



"9-7"
biorthogonal
easy to invert
smoother



extended Haar
biorthogonal
hard to invert
smooth (regular)

Notice, only 1-D filtering used.

Applications: Denoising

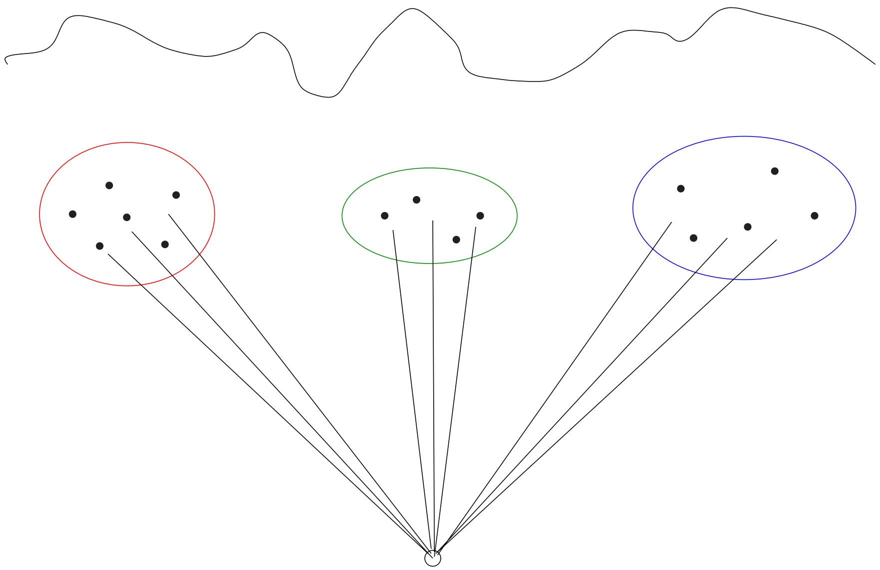
Original

10.66dB

2-D UWT 24.75dB

MDir UWT 25.94dB

Back to the Future:
Plenoptic Function and Sensor Networks



- The source is distributed in space
- Sensors observe correlated data
- In our context sensors are digital cameras
- We want to perform compression but we want to avoid communication among sensors.

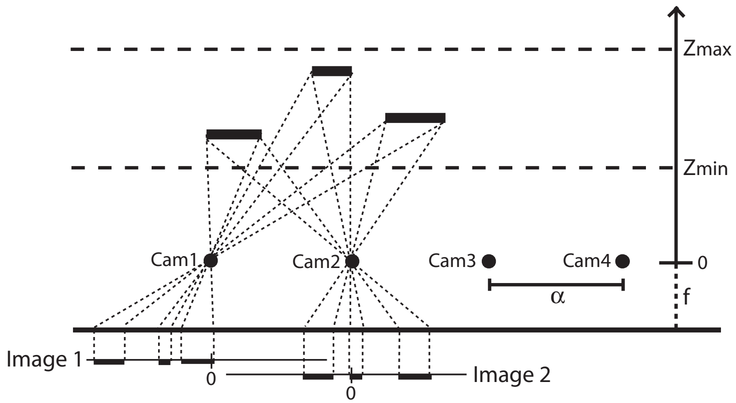
Background: Distributed Source Coding

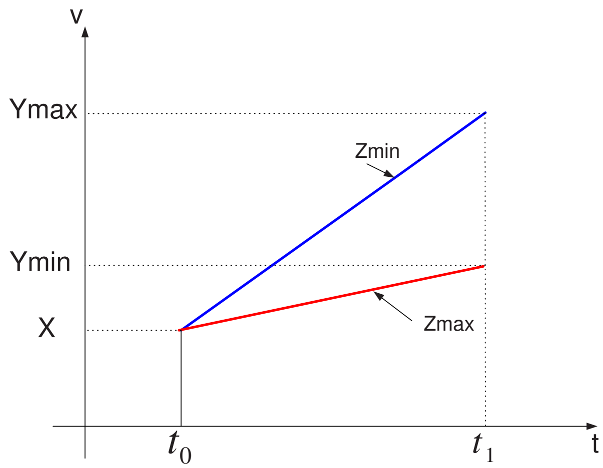
- Fundamental theoretical results
  - Slepian-Wolf 1973. Lossless coding
  - Wyner and Ziv 1976. Lossy coding with side information at the decoder.

- Constructive codes:
  - DISCUS [Pradhan-Ramchandran:99].
  - Extensions using advanced channels coded [Garcia-Frias:01, Aaron-Girod:02, Liveris-Xiong-Georghiades:02]

Plenoptic Constraints

In order to develop distributed compression algorithms, we need to know the correlation structure of the data. In the case of camera sensors, this structure is given by the Plenoptic function.

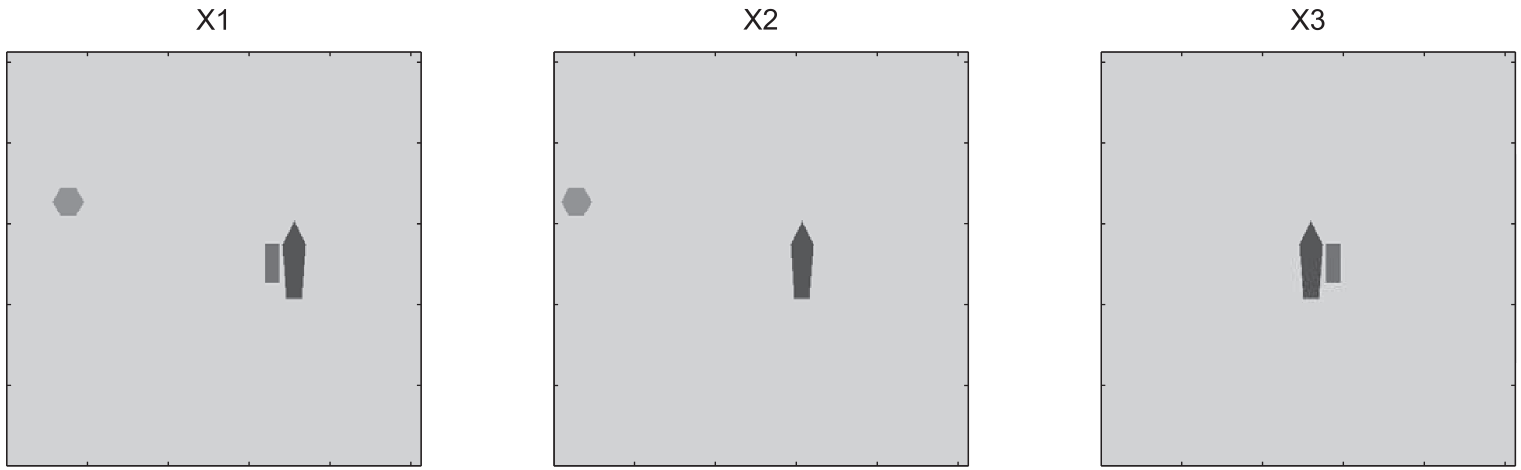




- Assume X and Y are the positions of a point on two different images
- Key insight: Whatever the complexity of the scene, if the depth of field is bounded, the disparity is also bounded.
- Use the information z_{min}, z_{max} and the camera locations, to develop distributed compression algorithms

Simulation Results

Three views of a simple synthetic scene



- Independent coding: 18bits per vertex
- Slepian-Wolf with no occlusions: 10bits per vertex
- Slepian-Wolf with occlusions: 14bits per vertex

Conclusions

What has the wavelet community given us?

- A unifying framework
- A new compression standard
- Improved modeling of images
- Good understanding of the interaction between representation, approximation and compression

But

- It is still not clear how to design efficient multi-dimensional bases,
- Interaction between continuous-time and discrete-time worlds
- What happens if we have only access to local information (Distributed Compression)?

"You **can't** predict the future, but you **can** invent it."
Dennis Gabor

Publications

http://www.commsp.ee.ic.ac.uk/∼pld/publications/

- On Wavelet Footprints
  - P.L. Dragotti and M. Vetterli, Wavelet Footprints: Theory, Algorithms and Applications, IEEE Trans. on Signal Processing, vol. 51(5), pp. 1306-1323, May 2003.
- On Quad-Tree Decompositions and Directional Wavelets
  - R. Shukla, P.L. Dragotti, M.N. Do and M. Vetterli, Rate-distortion optimized tree structured compression algorithms for piecewise polynomial images, IEEE Trans. on Image Processing, vol. 14 (3), pp. 343-359, March 2005.
  - V. Velisavljevic, B. Beferull-Lozano, M. Vetterli and P. L. Dragotti, Discrete Multi-Directional Wavelet Bases, Proc. of IEEE International Conference on Image Processing (ICIP), Barcelona, Spain, September 2003.
- On Distributed Source Coding
  - N. Gehrig and P.L. Dragotti, DIFFERENT DIstributed and Fully Flexible image EncodeRs for camEra sensor NeTworks, to be presented at the IEEE International Conference on Image Processing (ICIP), Genova, Italy, September 2005.
  - N. Gehrig and P.L. Dragotti, Distributed Compression of the Plenoptic Function, in Proc. of IEEE International Conference on Image Processing (ICIP), Singapore, October 2004.
  - N. Gehrig and P.L. Dragotti, Symmetric and Asymmetric Slepian-Wolf codes with systematic and non-systematic linear codes, IEEE Communications Letters, vol.9, pp. 61-63, January 2005.