Image Coding:
In search of Efficient Representations

François G. Meyer
Department of Electrical Engineering
University of Colorado at Boulder

francois.meyer@colorado.edu
http://ece-www.colorado.edu/∼fmeyer

Collaborators

- Amir Averbuch
- Raphy Coifman
- Jan-Olov Strömberg

- Image compression fundamentals
- Image coding with wavelets
- Coding of oscillatory texture

Image compression fundamentals

Image compression

What is the problem ?

1. approximation of an image for a given budget, or a given quality,
2. efficient computation : fast algorithms

Data compression techniques can be used for :

1. Discovering the structure of the data
2. Dimensionality reduction
3. Classification

Image compression: why ?

Applications	horiz $\times$ vert	frame/s	size	ratio
Digital camera	2560×1920		14 MB	10
Fax	1728×1100		240 kB	20
CD	352×240	30	7.6 Mb/s	50
HDTV	1920×1080	30	186 Mb/s	80
VideoPhone	176×144	15	1.1 Mb/s	300

Digital images files are large...

Important parameters of a compression system

- Compression efficiency: bit per pixel
- Fidelity: PSNR, visual inspection
- Complexity
- Robustness

Transform coding

- Apply a linear transformation to the image $f = \sum_{n=0}^{N-1} f_n \psi_n$
- A small number of coefficients f_n carry most of the energy
- Quantization of the coefficients : $\mathbb{R} \to \{1, \cdots, Q\}$
- Entropy coding: code $(q) \in \{0,1\}^*$, length (code): minimum

Key idea : At low bit rates, the distortion depends on the ability of the basis to approximate the signal with a small number of vectors [Mallat and Falzon, 1998]

$$\min_{\alpha_n} \|f - \sum_{n=0}^{M} \alpha_n \psi_n\| \; ; \quad M \ll N \tag{1}$$

Transform coding: Karhunen-Loève transform

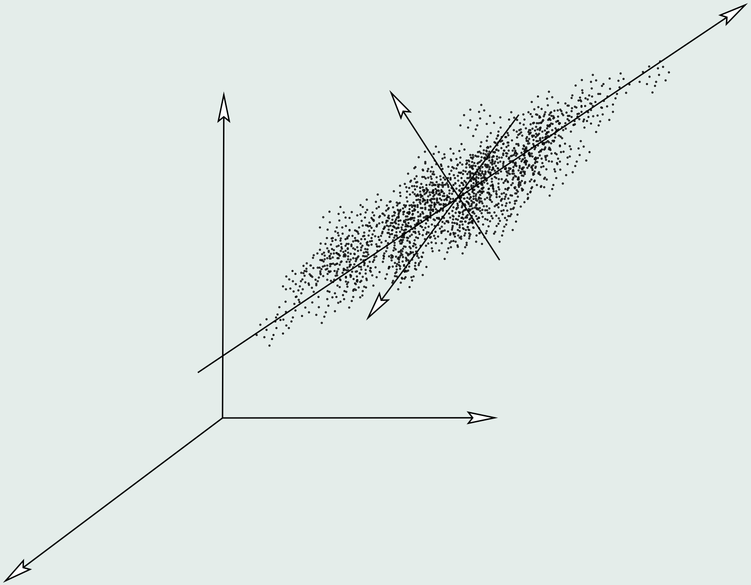
- $\mathbf{X} = (X_0, \cdots, X_{N-1})$ stochastic process, $\mathbf{R}_{\mathbf{X}}(i,j) = E[\mathbf{X}_i \mathbf{X}_j^T]$
- $(\boldsymbol{\mu}_0, \cdots, \boldsymbol{\mu}_{N-1})$ eigenvectors, $(\lambda_0, \cdots, \lambda_{N-1})$ eigenvalues of $\mathbf{R}_{\mathbf{X}}$.
- KLT : orthogonal transform $\mathbf{K} = \left[\boldsymbol{\mu}_0 | \cdots | \boldsymbol{\mu}_{N-1}\right]^T$
- de-correlate the image values :

$$\mathbf{Y} = \mathbf{K}\mathbf{X} \tag{2}$$

- If we keep only the first $M < N$ largest coefficients of $\mathbf{Y}$, then the KLT is the optimal orthogonal transform that minimizes the MSE.
- $\mathbf{X}$ Gaussian, high resolution scalar quantization, then KLT is optimal

$$\frac{D_{KL}(\bar{R})}{D_{\text{no transform}}(\bar{R})} = \frac{\left(\prod_{i=0}^{N-1} \lambda_i\right)^{1/N}}{\left(\sum_{i=10}^{N-1} \lambda_i\right)/N} \leq 1 \tag{3}$$

Karhunen-Loève transform



Some famous Karhunen-Loève transforms

- $\mathbf{X}$ stationary stochastic process$\rightarrow$ Discrete Fourier transform

$$\boldsymbol{\mu}_k = \frac{1}{\sqrt{N}} \left[1, \cdots, e^{\frac{2\pi i k n}{N}}, \cdots \right]^T \tag{4}$$

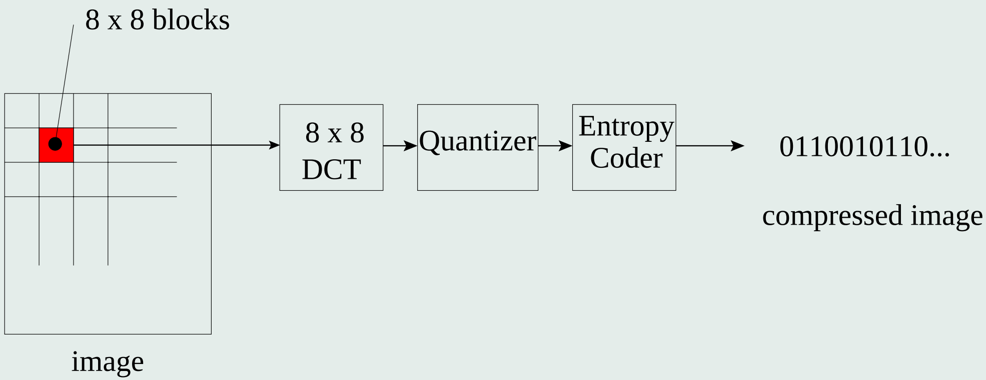
- $\mathbf{X}$ first order Gaussian Markov Process with high correlation $(\rho \rightarrow 1)$

$$\boldsymbol{\mu}_k = \frac{2}{\sqrt{N}} \left[\sqrt{2}, \cdots, \cos\left[\frac{(2n+1)k\pi}{2N} \right] \cdots \right] \tag{5}$$

  diagonalizes approximatively $\mathbf{R}_{\mathbf{X}}$
  $\rightarrow$ Discrete cosine transform

- In general : KLT not practical

JPEG Picture compression (1988)



Limitations :

- size of the blocks cannot be adapted to the content of the image
- no correlation between adjacent blocks: blocking effects

JPEG compression

Ratio = 64, PSNR=22.17dB

ratio = 62, PSNR=26.20dB

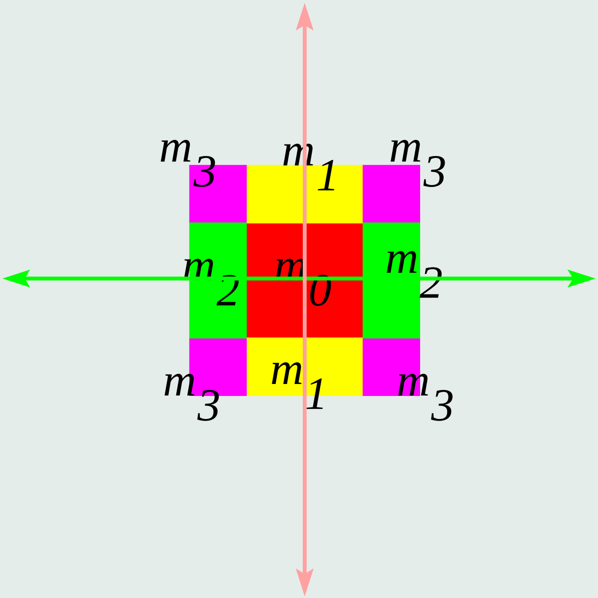
Image coding with wavelets
(from Bernouilli to Strömberg)

Are natural images scale invariant ?

Studies of large ensemble of natural images [Huang and Mumford, 1999, Ruderman and Blalek, 1994]:

- power spectrum (Fourier transform) $\sim C\xi^{-2}$
- I^k=average of the intensity over blocks of size $k \times k$:
  $D_H^k(i,j) = I^k(i,j+1) - I^k(i,j)$ does not depend on k
- suggest a self-similar process :
  if $X(\alpha t) = \alpha^H X(t)$ then $\Gamma_X(\xi) = C|\xi|^{-2H-1}$
- Should we use a wavelet transform for image coding ?

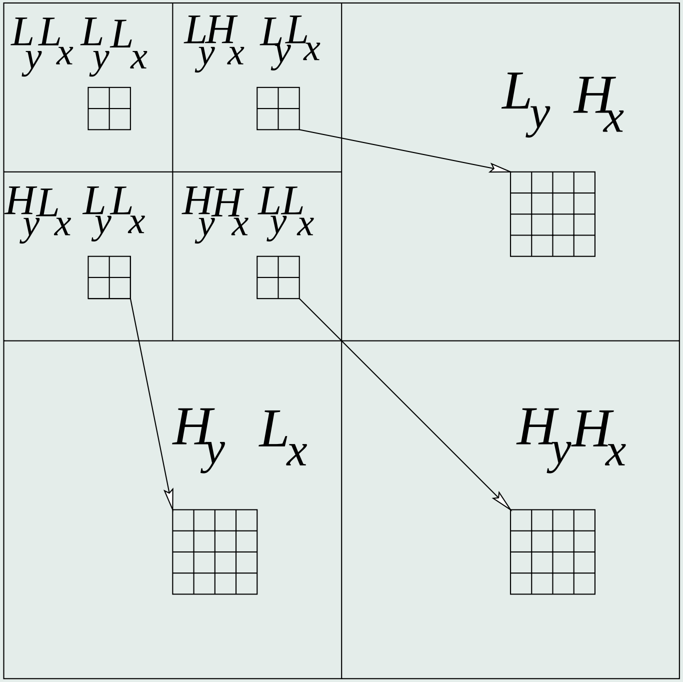
Two dimensional wavelet transform



Wavelet filters

one level wavelet decomposition

Wavelet image compression: zero tree coding [Shapiro, 1993, Said and Pearlman, 1996]



Wavelet transform: experimental observations

- clustering of coefficients at a given scale
- small and large coefficients at a given scale propagate at a fine scale:
  - cartoon model [Mumford, 1994]:
    image= smooth regions + edges
  - experimental findings from image ensembles:
    coefficients have similar statistics at all scale

Wavelet image compression: zero tree coding [Shapiro, 1993, Said and Pearlman, 1996]

Three symbols to characterize the quantized coefficients

1. ZTR: root of a zerotree: all children are quantized to zero,
2. POS: significant positive
3. NEG: significant negative
4. IZ, isolated zero: the coefficient is quantized to zero, but there exists some nonzero offspring

ZTR: codes zero jointly (vector quantization)

Details: choice of the filters

- linear phase vs orthogonality
- Size vs out of band rejection
- smoothness
- vanishing moments

Wavelet based compression: asymptotia ?

- Fast algorithm : $\mathcal{O}(N)$
- Very good quality for piecewise smooth images
- JPEG 2000
- BUT:
  - imprecise for high frequencies,
  - not adapted to texture (oscillatory patterns)

Coding of oscillatory texture

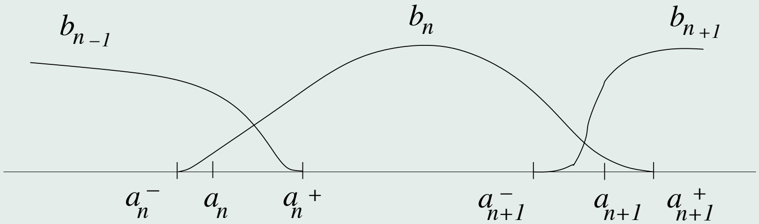
Representation of textured patterns

- simplest model of a patch of periodic texture :

$$w(x - x_0, y - y_0)\, e^{i(\xi x + \eta y)} \tag{6}$$

- local Fourier basis : most appropriate tool
- Can we fix JPEG :
  - blocks overlap smoothly
  - variable block size : adapted to the image content
  - fast algorithm to compute the optimal segmentation

Adaptive smooth local cosine transforms



$$\forall x \in [a_n^-, a_n^+], \quad b_n(x)\, b_{n-1}(2a_n - x) + b_n(2a_n - x)\, b_{n-1}(x) \neq 0 \tag{7}$$

$$\forall x \in [a_n^+, a_{n+1}^-] \quad b_n(x) \neq 0 \tag{8}$$

b_n is obtained from a prototype bell b

$$b_n(x) = b\left(\frac{x - a_n}{l}\right) \tag{9}$$

Dual bells

Let

$$\theta_n(x) = \frac{1}{b_n(x)\, b_{n-1}(2a_n - x) + b_n(2a_n - x)\, b_{n-1}(x)} \tag{10}$$

then the dual bell $\tilde{b}_n$ is defined as follows:

$$\tilde{b}_n(x) = \begin{cases} \theta_n(x)\, b_{n-1}(2a_n - x) & \text{if } a_n^- \leq x \leq a_n^+ \\ \dfrac{1}{b_n(x)} & \text{if } a_n^+ \leq x \leq a_{n+1}^- \\ \theta_{n+1}(x)\, b_{n+1}(2a_{n+1} - x) & \text{if } a_{n+1}^- \leq x \leq a_{n+1}^+ \\ 0 & \text{otherwise} \end{cases} \tag{11}$$

Biorthogonal local cosine bases

- basis functions of the DCT-IV:

$$C_{n,k}(x) = \sqrt{\frac{2}{a_{n+1} - a_n}} \cos\left[(k + 1/2)\pi \frac{x - a_n}{a_{n+1} - a_n}\right] \tag{12}$$

- Local cosine basis functions :

$$\psi_{n,k}(x) = b_n(x)\, C_{n,k}(x) \tag{13}$$

- dual basis functions :

$$\widetilde{\psi}_{n,k}(x) = \tilde{b}_n(x)\, C_{n,k}(x). \tag{14}$$

- $\exists B > A > 0$ such that,

$$A \sum_{n,k} |f_{n,k}|^2 \leq \| \sum_{n,k} f_{n,k} \psi_{n,k} \|^2 \leq B \sum_{n,k} |f_{n,k}|^2 \tag{15}$$

Choice of the bell function

- good approximation of polynomials
- reproducing $p^0(x) \equiv 1$ with exactly one coefficient $p^0_{k_0}$ per interval:

$$\forall x \in \mathbb{R}, \quad p^0(x) = \sum_n p^0_{k_0} \widetilde{\psi}_{n,k_0}(x)\ ; \quad p^0_{n,k} = p^0_{k_0}\, \delta_{k,k_0} \tag{16}$$

- unique symmetric solution

$$b(x) = \sin \frac{\pi}{2}(x + 1/2) \tag{17}$$

...but bell is not differentiable at $x = -1/2$ and $3/2$!

Optimized bells of Gregory Matviyenko [Matviyenko, 1996]

- good approximation of $p^0 \equiv 1$ over $[a_n, a_{n+1})$ using the first K coefficients $p^0_{n,k}, k = 0, \dots, K-1$.

$$\sum_{k=0}^{K-1} p^0_{n,k} \widetilde{\psi}_{n,k} \qquad \text{with} \quad p^0_{n,k} = \int p^0(x)\, \psi_{n,k}(x)\, dx \tag{18}$$

- find the bell b that minimizes the residual error :

$$\min_b \sum_{k=K}^{\infty} |p^0_{n,k}|^2 \tag{19}$$

under the constraint : $b(x) + b(-x) = 1 \quad$ for all $x \in [0, 1/2]$.

- Solution :

$$b^K(x) = \begin{cases} \frac{1}{2}(1 + \sum_{k=0}^{K-1} g_k \sin(k+\frac{1}{2})\pi x) & \text{if } x \in [-\frac{1}{2}, \frac{1}{2}) \\ \frac{1}{2}(1 + \sum_{k=0}^{K-1} (-1)^k g_k \cos(k+\frac{1}{2})\pi x) & \text{if } x \in [\frac{1}{2}, \frac{3}{2}) \\ 0 & \text{otherwise} \end{cases} \quad (20)$$

g_k are calculated numerically in [Matviyenko, 1996].

- $0 \le |b_K(x)| \le 1 \qquad 0 \le |\tilde{b}_K(x)| \le (\sqrt{2}+1)/2$
- Riesz bounds : $A = 1$ and $B = 2$

Modulated Lapped Biorthogonal Transform [Malvar, 1998]

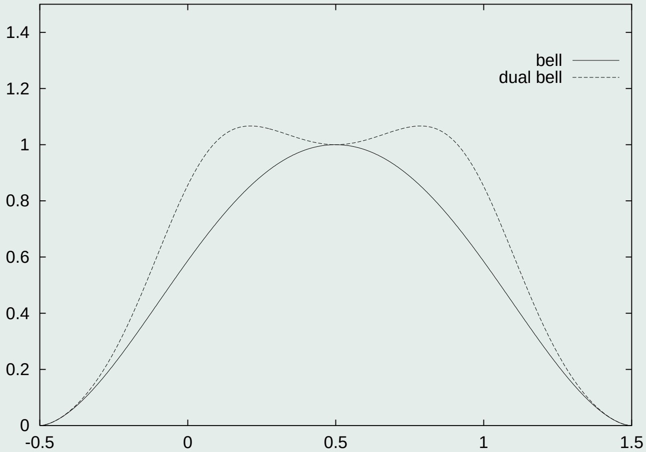
Idea : smooth $\sin \frac{\pi}{2}(x+1/2)$ at both end-points : take the square.

$$b(x) = \left[\sin \frac{\pi}{2}(x+1/2)\right]^2 = \frac{1-\cos \pi(x+1/2)}{2} \tag{21}$$

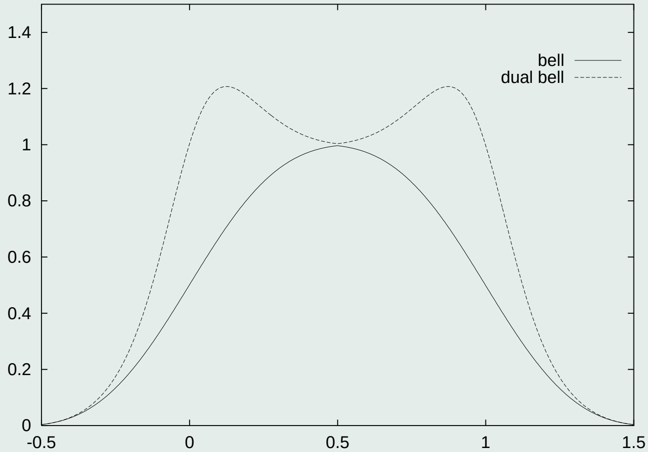
- MLBT :

$$b(x) = \begin{cases} \dfrac{1-\cos[\pi(x+1/2)^\alpha]+\beta}{2+\beta} & \text{if } x \in [-\frac{1}{2}, \frac{1}{2}] \\ b(x) = b(1-x) & \text{if } x \in [\frac{1}{2}, \frac{3}{2}] \end{cases} \tag{22}$$

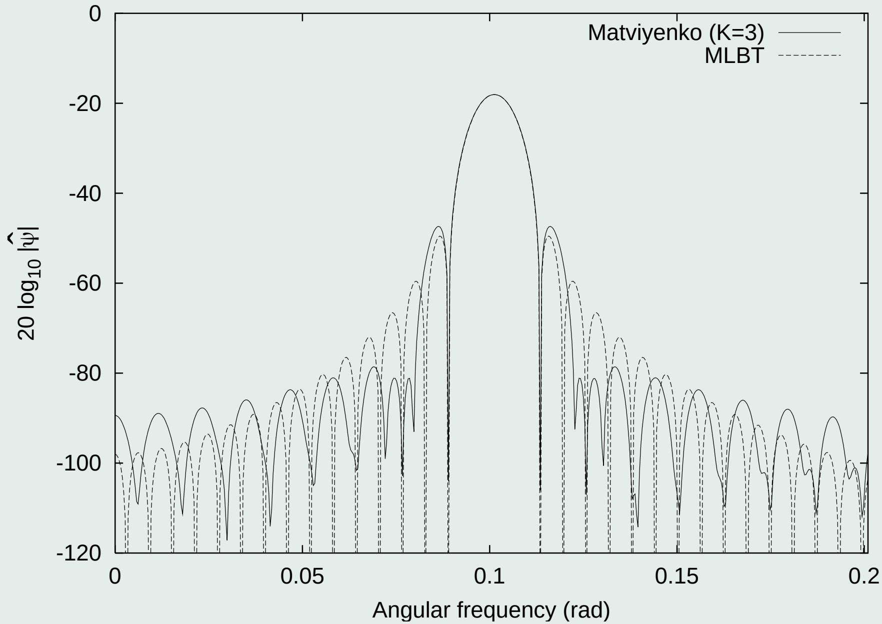
- if $\alpha > 1, \beta = 0$ then $b \in C^1$



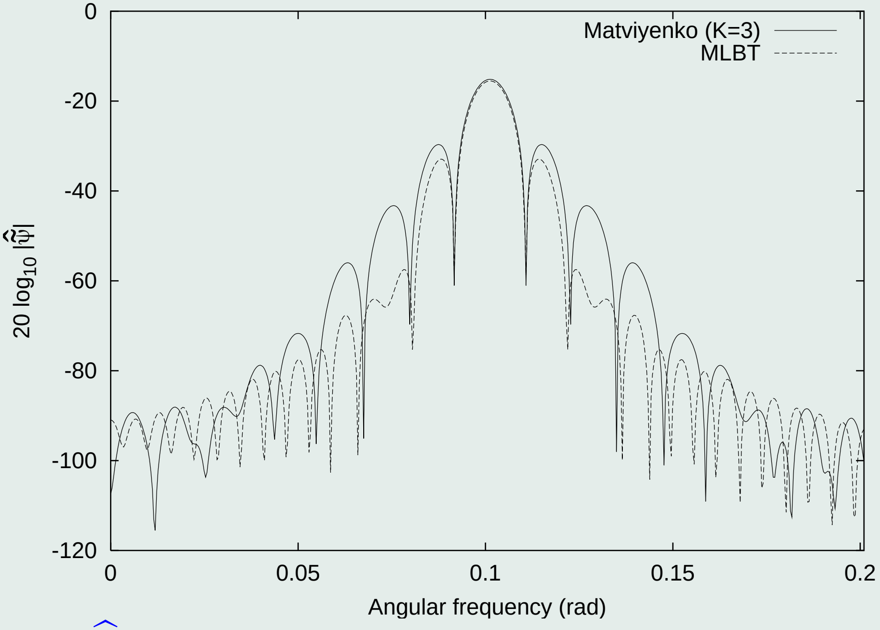
Matviyenko $b^3, \tilde{b^3}$.



MLBT's bell $b, \tilde{b}$



$20\log_{10}|\hat{\psi}_{n,16}|$ for Matviyenko's bell ($K = 3$), and the MLBT bell.

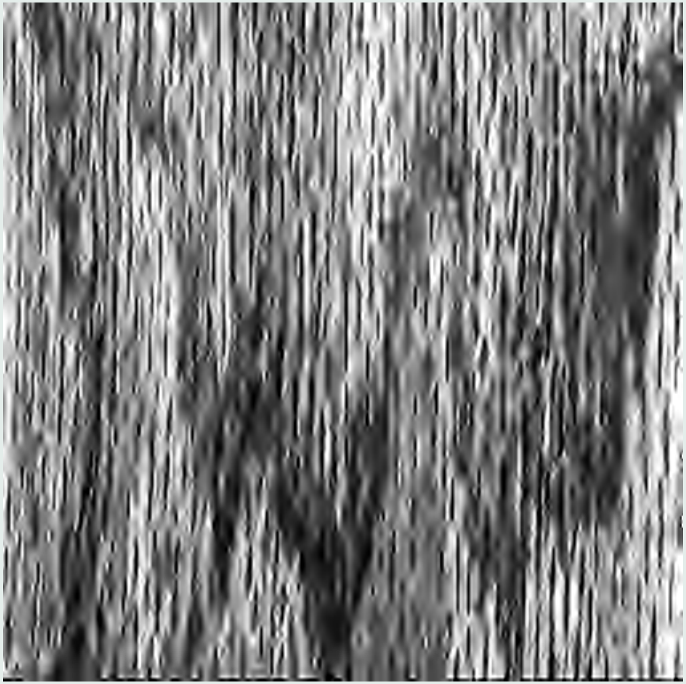


$20 \log_{10} |\widehat{\widetilde{\psi}}_{n,16}|$ for Matviyenko's bell, and the MLBT bell.

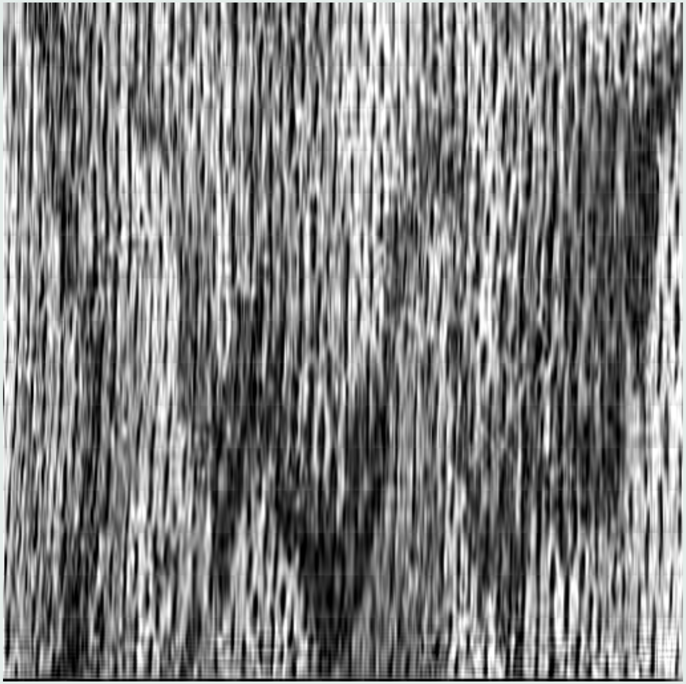
LCT compression algorithm [Meyer, 2002]

- best basis : quadtree segmentation
- cost function: estimate of the actual rate achieved by each node
- near optimal Laplacian scalar quantizer [Sullivan, 1996]
- ordering of the coefficients : large scale correlation between blocks
- bit plane encoding

Wood grain at 0.125 bpp (compression = 64).



SPIHT, PSNR=16.49dB



Matviyenko's bell, PSNR = 18.55dB.

Clown at 0.125bp (compression =64)

SPIHT, PSNR = 28.23 dB

MLBT bell, PSNR = 27.43dB.

Roof at 0.125 bpp (compression = 64).

SPIHT, PSNR=23.77 dB

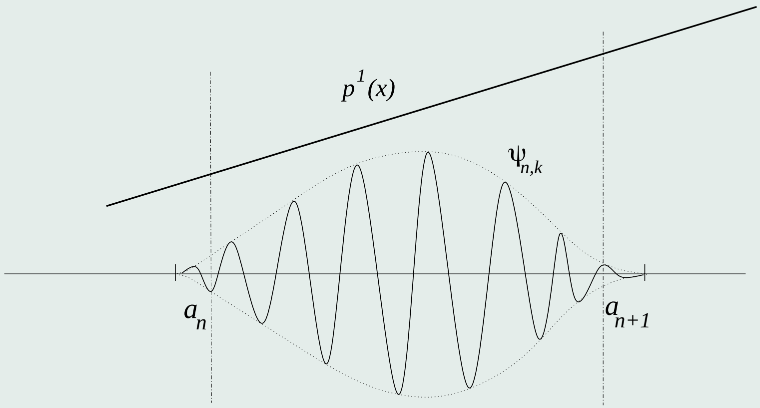
MLBT bell, PSNR = 24.55dB.

Optimized bells of Kai Bittner [Bittner, 1999]

- Idea : reproduce exactly polynomials with K coefficients, BUT non symmetric bells

- only possible with polynomial of degree ≤ 1

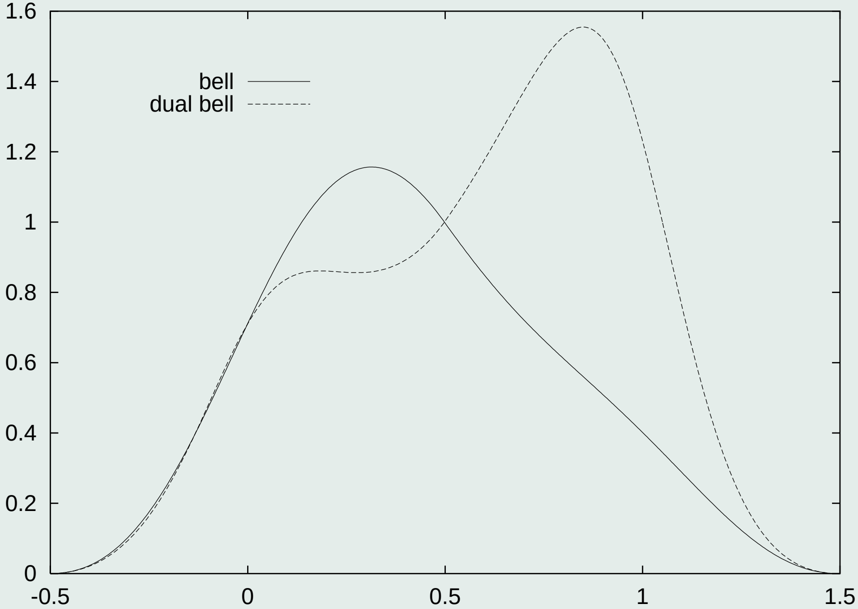
$$\text{if} \quad p^1(x) \equiv x \quad \text{then} \quad p^1_{n,k} =< p^1, \psi_{n,k} >= 0 \quad \text{for } k \geq K. \tag{23}$$



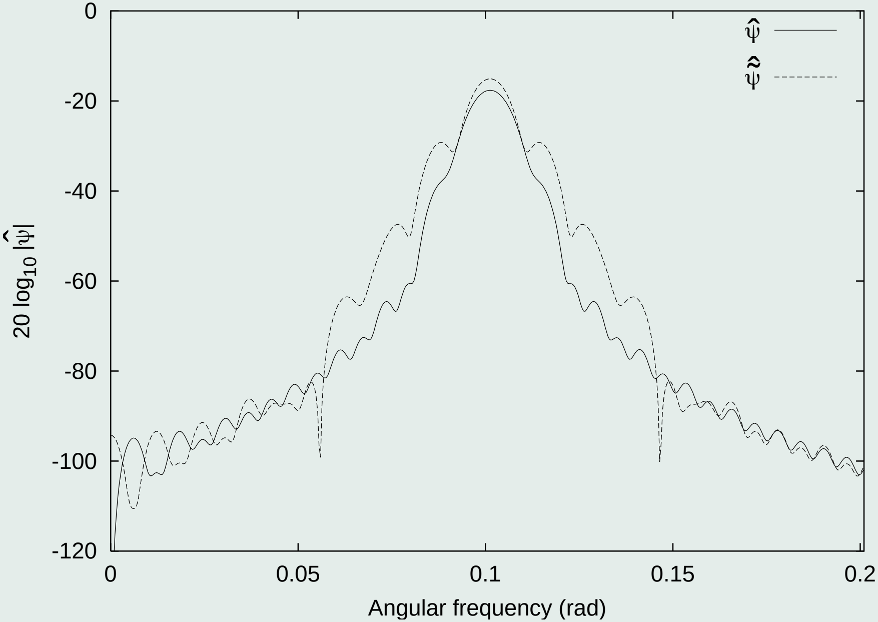
Optimized bells of Kai Bittner [Bittner, 1999]

- one unique solution to (23) with $b \in C^{K-1}$, and $\text{supp}(b) \subset [-\frac{1}{2}, \frac{3}{2}]$.
- $\{\psi_{n,k}\}$ forms a Riesz basis $\iff$ K is odd
- Riesz bound B increases very rapidly with K

- little practical use for compression if $K \geq 5$
- We use $K = 3$, Riesz bounds : $A = 0.742$ and $B = 3.067$



Bittner's bell optimized for linear functions : graphs of b and $\tilde{b}$.

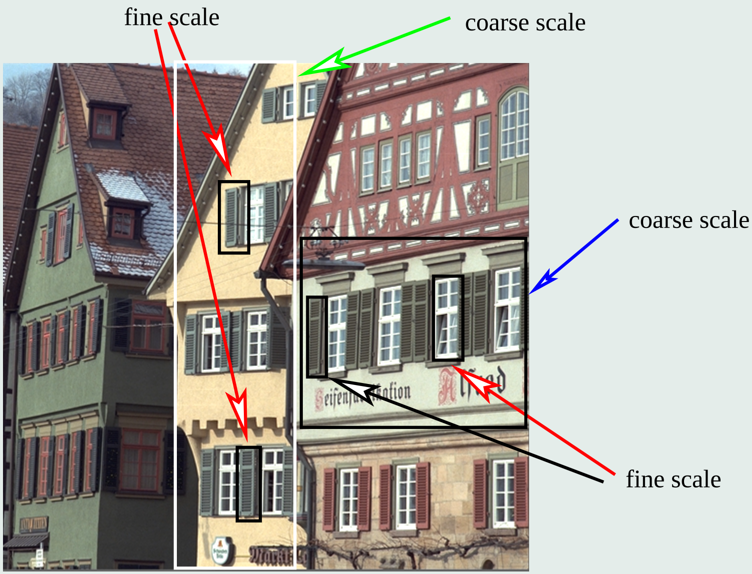


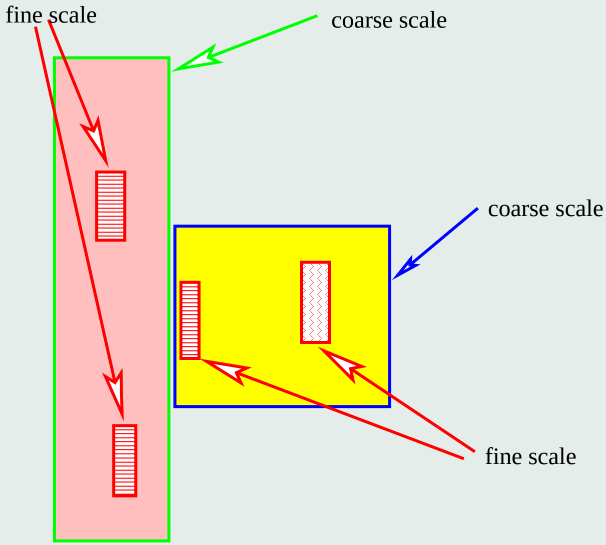
Bittner's bell : magnitude of the Fourier transform of $\psi_{n,16}$, and $\widetilde{\psi}_{n,16}$.

Roof at 0.125 bpp (compression = 64).

Bittner's bell, PSNR=24.12 dB.

Limitations of local Fourier approaches





- replace the block segmentation by a superposition of filtered versions of the original image
- replace the octave band decomposition by a more general splitting of the Fourier domain.

Fast Wavelet Packet image compression [Meyer et al., 2000]

1. fast 2-D convolution-decimation algorithm with factorized non-separable 2-D filters : 4 times faster (no transpose)
2. cost function that takes into account the cost of coding the output levels of the quantizers, and the cost of coding the significance map
3. context-based entropy coder that uses a space filling curve.

Roof at 0.125 bpp (compression = 64).

SPIHT, PSNR=23.77 dB

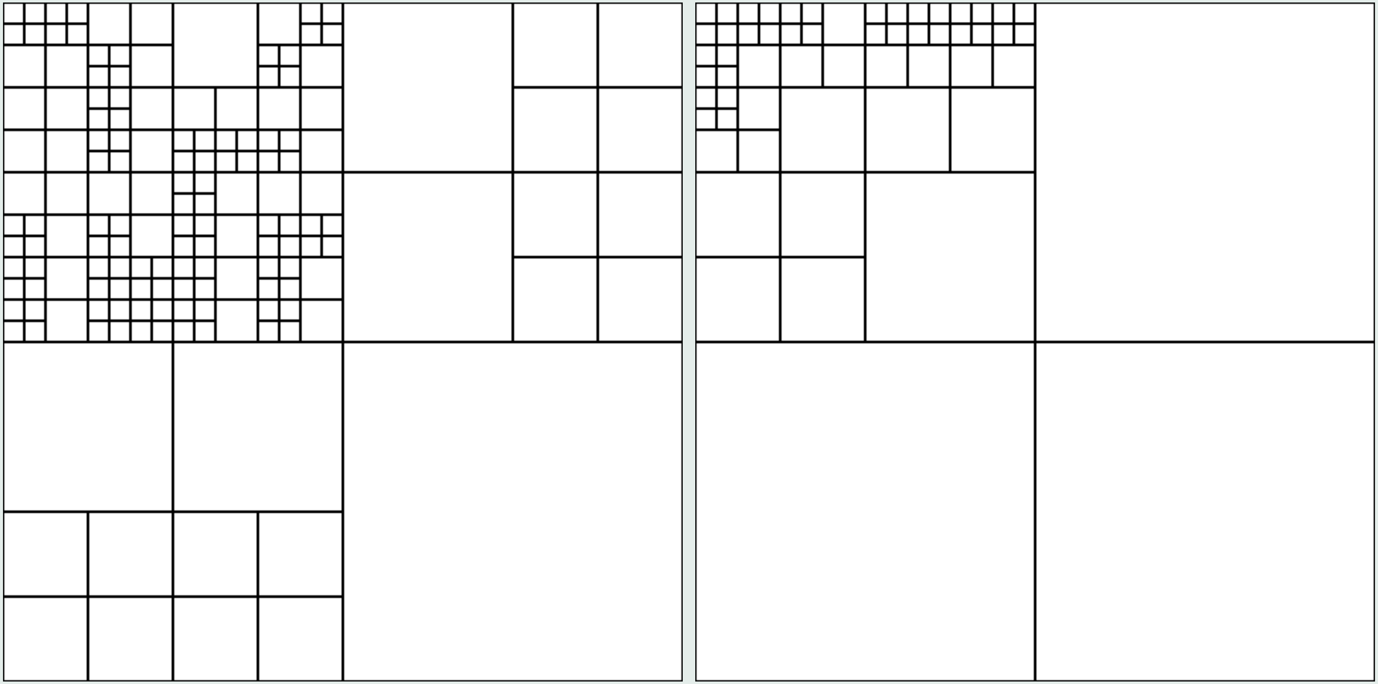
FWP, PSNR=25.10 dB

Clown at 0.125 bpp (compression = 64).

SPIHT, PSNR=28.23dB.

FWP, PSNR=28.49 dB

Best basis for roof and clown at 0.125 bpp (compression = 64)



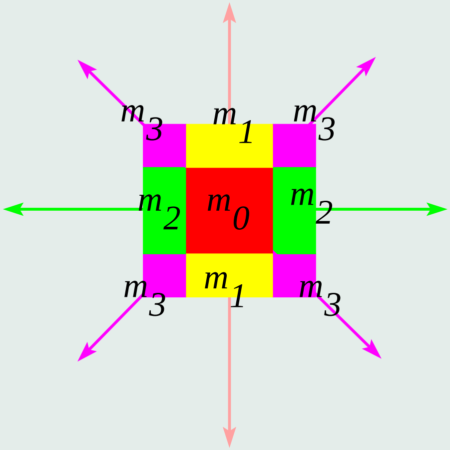
Roof

Clown

- textures : oriented with all possible directions, frequencies, and locations.
- tensor product of wavelets :

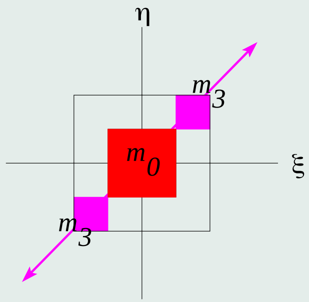
$$\psi_k(x,y) = \begin{cases} \varphi(x)\varphi(y) & \text{if } k = 0 \\ \varphi(x)\psi(y) & \text{if } k = 1 \\ \psi(x)\varphi(y) & \text{if } k = 2 \\ \psi(x)\psi(y) & \text{if } k = 3 \end{cases} \tag{24}$$

- wavelet filter banks m_1, m_2, m_3 can resolve 2.5 directions



- Wavelet packets : more directions...
- BUT Fourier transform of the tensor product of two real valued wavelet packets : four symmetric peaks in the frequency plane.

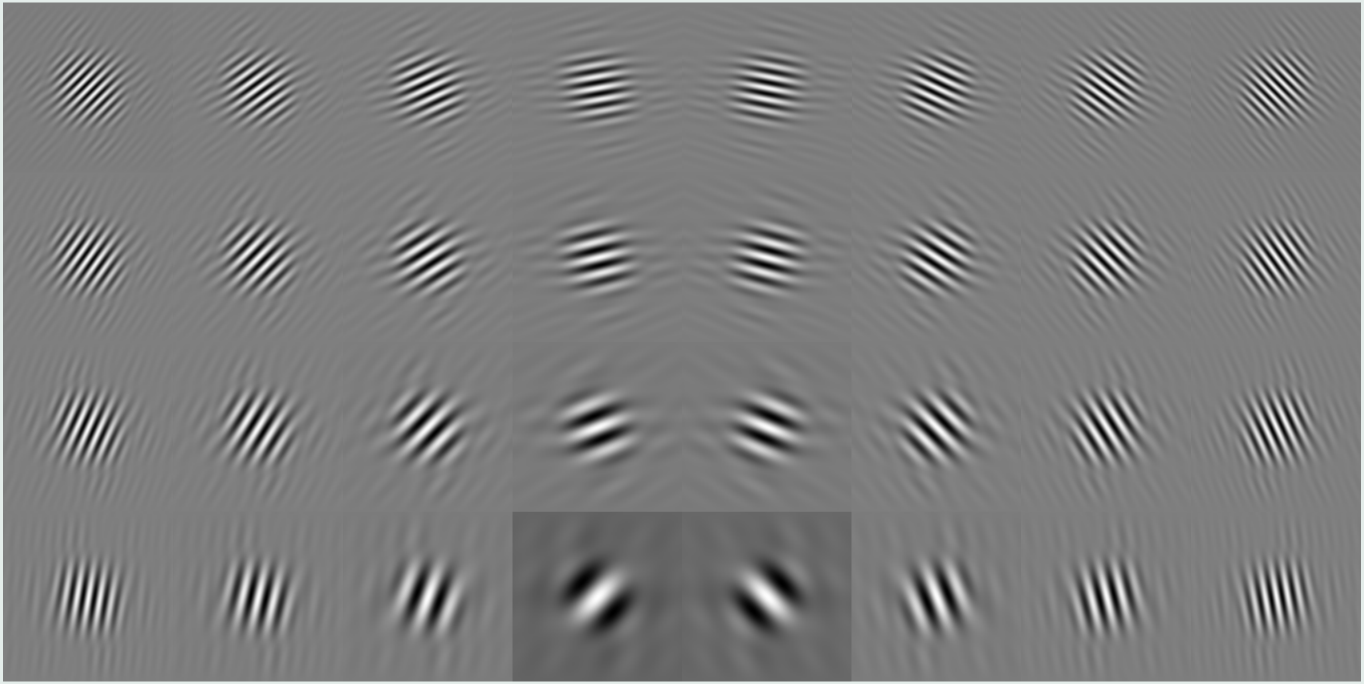
- geometric interpretation of a large wavelet packet coefficient is problematic :
  - the intensity is oscillating as a planar wave $e^{i(\omega_x x + y\omega_y y)}$,
  OR
  - the intensity is oscillating with the conjugate frequency $e^{i(\omega_x x - y\omega_y y)}$.
- remove the ghost in the conjugate direction :
  steerable wavelet packets



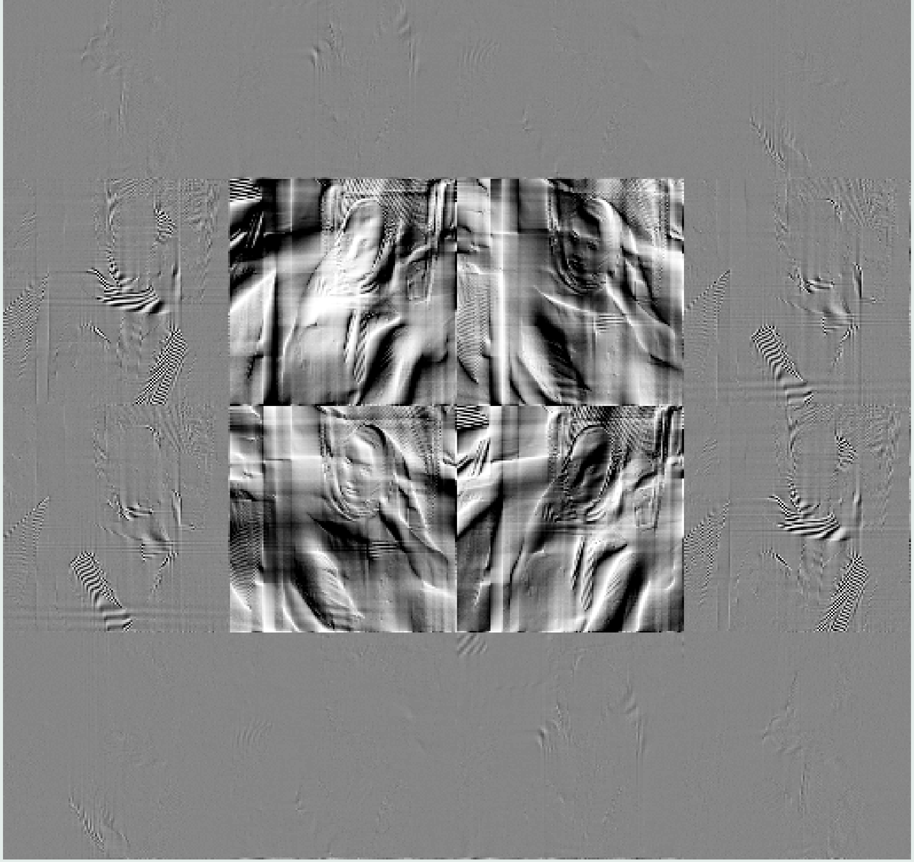
Brushlets [Meyer and Coifman, 1997, Meyer et al., 2002]

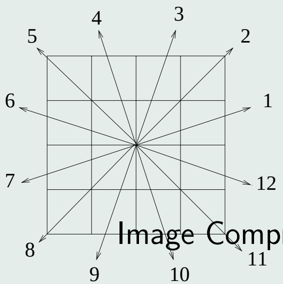
- duality between local trigonometric bases and wavelet packets
- construction : expansion of the Fourier transform of the image into local Fourier bases
- brushlets are complex valued functions with a phase

Brushlets



Basis functions $\psi_{m,j} \otimes \psi_{n,k}$ for the frequencies $(m\frac{2\pi}{512}, n\frac{2\pi}{512})$, with $(m,n) \in \{-48,-32,-16,0,16,32,48\}^2$. We have $h_m = l_n = 16, \delta = 8$



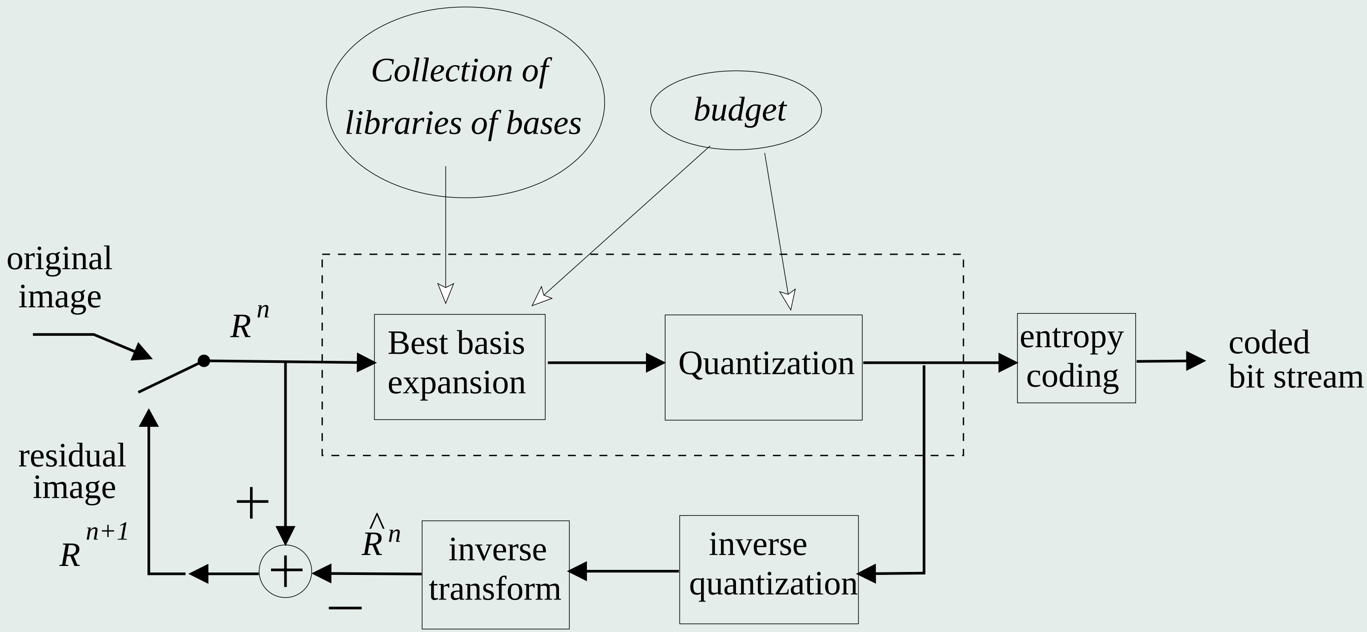

4
3
5
2
6
1
7
12
8
11
9
10

Multi-layer image representation [Meyer et al., 2002]

image = sum of two layers : a “cartoon image” and a texture map.

- cartoon = salient parts, piecewise smooth changes in the illumination
- texture map = texture in the regions enclosed by edges
- cartoon and texture map should be represented with two different sets of basis functions
- algorithm : cascade of compressions applied successively to the image itself and to the residuals from the previous compressions
- code the residual part in a lossy way: retain only the most significant structures of the residual part

Multi-layer image representation



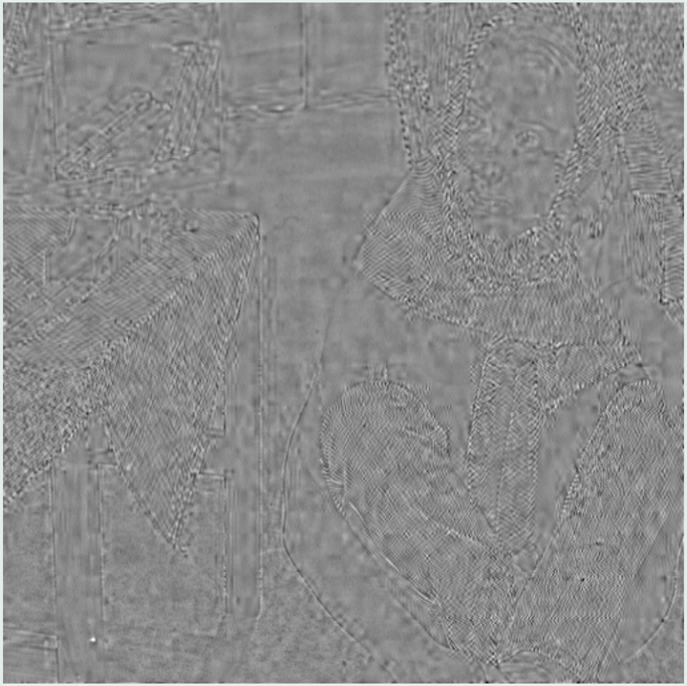
A toy example

compression by 254 with wavelets.

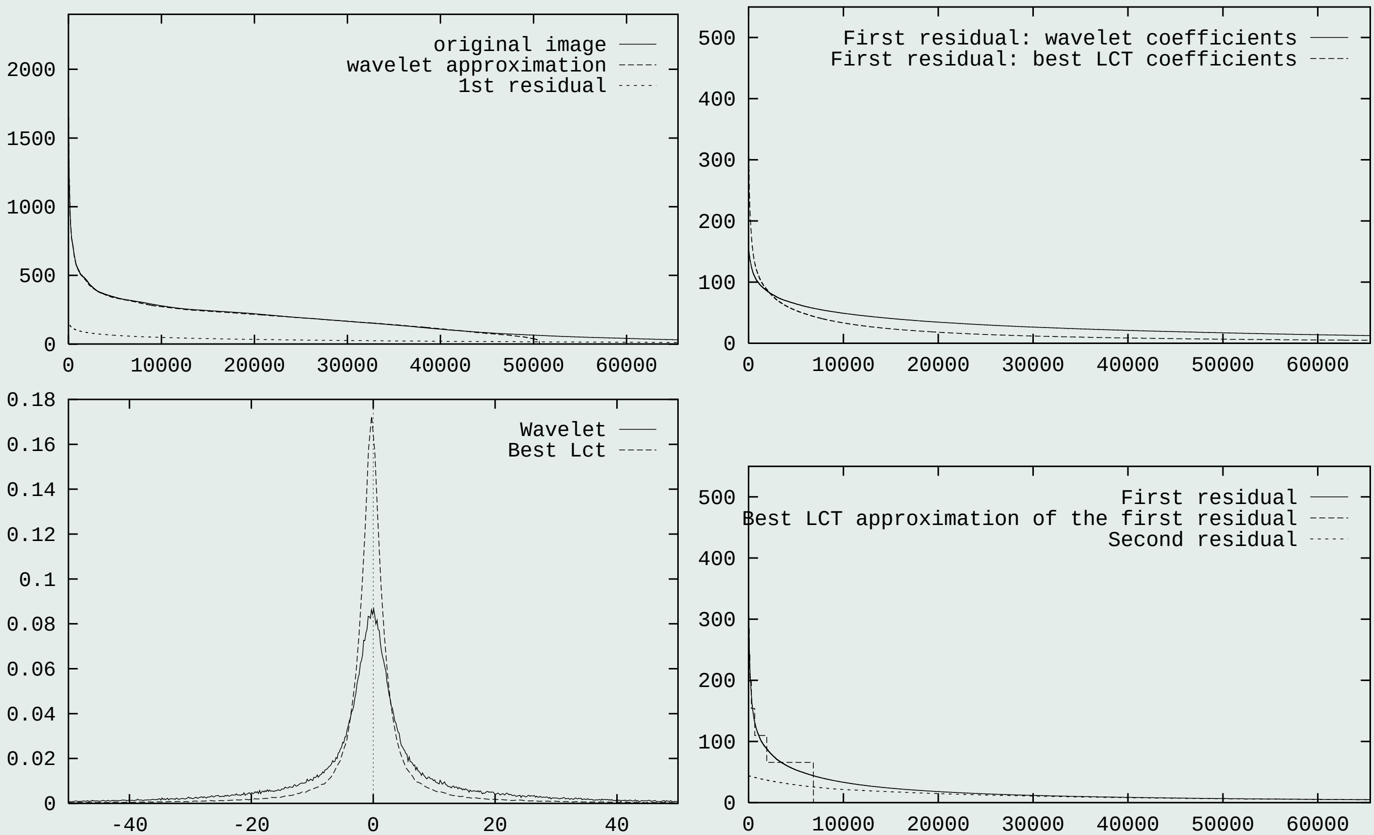
first residual.

A toy example (2)

compression by 36 (LCT)



Second residual.


original image
wavelet approximation
1st residual
0 500 1000 1500 2000
0 10000 20000 30000 40000 50000 60000
First residual: wavelet coefficients
First residual: best LCT coefficients
0 100 200 300 400 500
0 10000 20000 30000 40000 50000 60000
Wavelet
Best Lct
0 0.02 0.04 0.06 0.08 0.1 0.12 0.14 0.16 0.18
-40 -20 0 20 40
First residual
Best LCT approximation of the first residual
Second residual
0 100 200 300 400 500
0 10000 20000 30000 40000 50000 60000

original image

compression by 150

compression by 10

final residual

References

[Bittner, 1999] Bittner, K. (1999). Error estimates and reproduction of polynomials for biorthogonal local trigonometric bases. *Applied and Computational Harmonic Analysis*, 6:75–102.

[Huang and Mumford, 1999] Huang, J. and Mumford, D. (1999). Statistics of natural images and models. In *Proc IEEE Conf. on Computer Vision and Pattern Recognition*, volume 1, pages 541–547. IEEE Comput. Soc.

[Mallat and Falzon, 1998] Mallat, S. and Falzon, F. (1998). Analysis of low bit rate image transform coding. *IEEE Trans. on Signal Processing*, 46(4):1027–42.

[Malvar, 1998] Malvar, H. (1998). Biorthogonal and nonuniform lapped transforms for transform coding. *IEEE Trans. Signal Processing*, 46(4):1043–53.

[Matviyenko, 1996] Matviyenko, G. (1996). Optimized local trigonometric bases. *Applied and Computational Harmonic Analysis*, 3:301–23.

[Meyer, 2002] Meyer, F. (2002). Image compression with adaptive local cosines : A comparative study. *IEEE Trans. Image Processing*, pages 616–629.

[Meyer et al., 2002] Meyer, F., Averbuch, A., and Coifman, R. (2002). Multi-layered image representation:application to image compression. *IEEE Transactions on Image Processing*, pages 1072–1080.

[Meyer et al., 2000] Meyer, F., Averbuch, A., and Strömberg, J.-O. (2000). Fast adaptive wavelet packet image compression. *IEEE Trans. Image Processing*, pages 792–800.

[Meyer and Coifman, 1997] Meyer, F. and Coifman, R. (1997). Brushlets: a tool for directional image analysis and image compression. *Applied and Computational Harmonic Analysis*, pages 147–87.

[Mumford, 1994] Mumford, D. (1994). The bayesian rationale for energy functionals. In *Geometry-Driven Diffusion in Computer Vision*, pages 141–153.

[Ruderman and Blalek, 1994] Ruderman, D. and Blalek, W. (1994). Statistics of natural images : Scaling in the woods. *Physical Review Letters*, 73(6):814–817.

[Said and Pearlman, 1996] Said, A. and Pearlman, W. (1996). A new fast and efficient image codec based on set partioning in hierarchical trees. *IEEE Trans. Circuits. Syst. Video Tech.*, 6:243–50.

[Shapiro, 1993] Shapiro, J. (1993). Embedded image coding using zerotrees of wavelet coefficients. *IEEE Trans. on Signal Processing*, 41(12):3445–3462.

[Sullivan, 1996] Sullivan, G. (1996). Efficient scalar quantization of exponential and Laplacian random variables. *IEEE Trans. Information Theory*, 42(5):1365–1374.