Techniques for Video Compression and Analysis (5LSE0), Module 04

Wavelet Transformation and the JPEG2000 standard

Peter H.N. de With
(p.h.n.de.with@tue.nl)

slides version 1.0

5LSE0 - Mod 04
Part 01

Introduction to Wavelets and the Discrete Wavelet Transform

Overview of Module 04

* **A1: Wavelet transform**
  – Introduction and Fourier analysis, extend to wavelet transform
  – Subband filters, Quadrature Mirror Filters
  – Wavelets, functions and banks
* **A2: Wavelet video coding**
  – Wavelet coefficient coding for compression
* **B: JPEG2000 standard**
  – Requirements, applications, standard principles
  – EBCOT coding, forms of scalability

Introduction / random signal

* **Any kind of signal contains of low and high fre-quencies, variations in amplitude, ... basis functions?**
* **For signal analysis one would like a tool, which provides these properties:**
  – Frequency
  – Amplitude
  – Location
* **Motivation: analysis components can be used as input for (a.o.) compression systems.**
  – Which “tools” suffice?

Fourier transform

Fourier transform based on a signal decomposition of **sine/cosine** waves:

$$F(\omega) = \int_{-\infty}^{\infty} f(t) e^{-j\omega t} dt$$

* **It has:**
  - Good sine wave detection/recognition
  - No time-localization property: averaging over time!
  - No transient detection
* **Good tool for frequency analysis where time-localization is not important.**

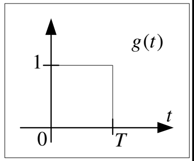
Short-Time Fourier Transform (STFT) –(1)

* **Because: if $f(t)$ is a non-periodic signal, the summation of periodic functions, sine & cosine (as with Fourier), does not accurately represent the signal.**
* **STFT: decompose the signal $f(t)$ into pieces to achieve time-localization by means of a "window"-function**

  (a.k.a. Windowed Fourier Transform)

Short-Time Fourier Transform – (2)

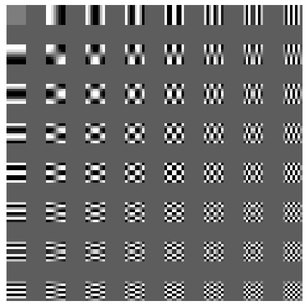
* **Short-Time:**
  - Definition: $F(\omega, \tau) = \int_{-\infty}^{\infty} f(t) g(t-\tau) e^{-j\omega t} dt$



  - Where *g(t)* is a rectangular window, *g(t)*=1 for *t*=<0,*T*] and *g(t)*=0 otherwise
  - Sometimes other window functions are used e.g. if *g(t)* is Gaussian, the STFT is the Gabor transform
  - Note: Able to analyze frequencies, able to do time-localization but…the localization window is fixed! Hereby no scaling along with the (low-high) frequencies.

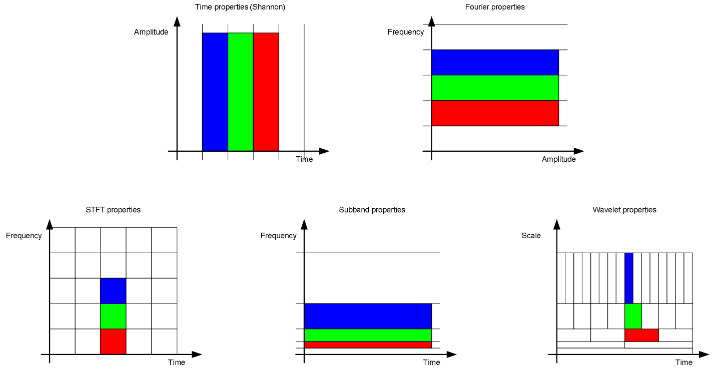
Short-Time Fourier Transform – (3)

* **Image & video coding is based on DCT**
  - 8x8 video block sampling
* **2D DCT decomposition**
  - for 8x8 blocks
  - Figure shows basis functions
  - Patterns for projections
* **However ,…**
  - Block aliasing occurs



Overview of time-frequency for different decompositions

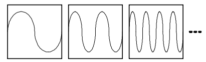
9



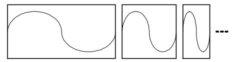
Wavelets / Introduction

10

* With the STFT, the time window remains equal while other cycles (i.e. frequency) of the *basis function* (sine/cosine) existed in that window



* With a *wavelet basis function*, the window size changes while the #cycles remains equal



* Cycle and location detection are done by "scaling" and "translating" the *basis function* / *mother wavelet*

TU/e EINDHOVEN UNIVERSITY OF TECHNOLOGY PdW-YSu / 2022-23 Fac EE / SPS-VCA 5LSE0 / Module 04 Wavelets, DWT and main appl. VCA

Wavelets / Mother wavelet

11

* 'Mother' maintains its properties, derive the remaining scaled and translated wavelet functions ("psi") by:

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}}\psi\left(\frac{t-b}{a}\right)$$

where ***a*** represents **scaling** and ***b*** represents **translation**

* **Scaling** $\frac{1}{\sqrt{a}}$ is needed to maintain "norm"

$$\|f(t)\|^2 = \int_{-\infty}^{\infty} f^2(t)dt$$

TU/e EINDHOVEN UNIVERSITY OF TECHNOLOGY PdW-YSu / 2022-23 Fac EE / SPS-VCA 5LSE0 / Module 04 Wavelets, DWT and main appl. VCA

Wavelets / Continuous functions

12

* Defining Continuous Wavelet Transform (CWT)

$$y(a,b) = \int_{-\infty}^{\infty} x(t)\cdot\psi_{a,b}(t)dt$$

* and its inverse

$$x(t) = \frac{1}{C_\psi}\int_{-\infty}^{\infty}\int_{0}^{\infty}\frac{1}{a^2}\,y(a,b)\,\psi_{a,b}(t)\,da\,db$$

where $C_\psi = \int_0^\infty \frac{|\Psi(\omega)|^2}{\omega}d\omega$ and $\Psi(\omega) = F[\psi(t)]$ as its Fourier transform

TU/e EINDHOVEN UNIVERSITY OF TECHNOLOGY PdW-YSu / 2022-23 Fac EE / SPS-VCA 5LSE0 / Module 04 Wavelets, DWT and main appl. VCA

Wavelets / Energy constraint

- Needed $\Psi(0)=0$ for the integral to exist: $\int_{-\infty}^{\infty}\psi(t)dt=0$
  - ➢Hence the wavelet function is a “specific” HP-filter, avg = 0
- And we would like finite energy (via Parseval)

$\int_{-\infty}^{\infty}|\Psi(\omega)|^2 d\omega<\infty$ this happens if $|\Psi(\omega)|^2$ decays when $\omega \rightarrow \infty$

Energy now in narrow frequency band: => frequency localization!

Wavelets / Definition of Discrete WT

- **Discrete Wavelet Transform (DWT)**

let $a=a_0^{-m},\quad b=n\cdot b_0\cdot a_0^{-m}$

then $\psi_{m,n}(t)=a_0^{m/2}\cdot\psi(a_0^m t-nb_0),\quad m,n\in Z$

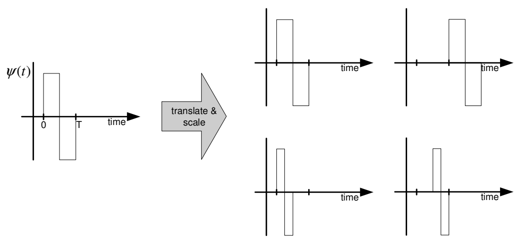
for $a_0=2,\quad b_0=1$, we have $\psi_{m,n}(t)=2^{m/2}\cdot\psi(2^m t-n)$

- ➢Dyadic wavelets: integer time shift and power of two scaling.
- ➢…these are however not the only possible DWT.

Wavelets / Example of shifting, scaling

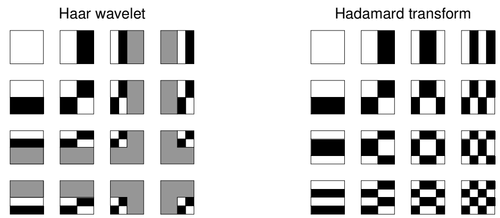
- **Example: the Haar wavelet**

$$\psi(t)=\begin{cases}1 & 0\le t<\frac{1}{2}\\ -1 & \frac{1}{2}\le t<1\end{cases}$$



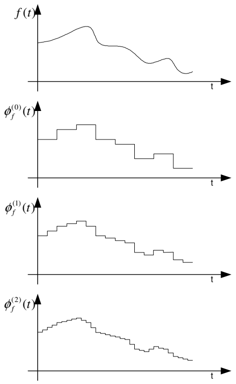
Wavelets / Compare Haar & Hadamard

- **Haar wavelet versus the Hadamard transform**
  - 2D decomposition (2x1D)



Wavelets / Scaling function

- **Scaling function** $\phi(t)$ **has the property that function** $f(t)$ **can be represented by the scaling function, but also be represented by dilated versions of the scaling function ("phi")**
- **Scaling function acts as the LP counterpart of the wavelet**
- **Hence: composition is phi+psi!**



Wavelets / Multi Resolution Analysis (MRA) – (1)

- **Scaling function can also be represented by its dilations at a higher resolution (special case with half-band filters)**

$$\phi(t) = \sum_k h_k \sqrt{2}\phi(2t-k)$$

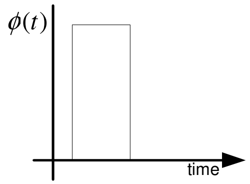
- **A.k.a. dilation equation**

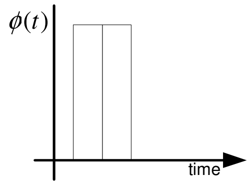
Wavelets / Multi Resolution Analysis (MRA) – (2)

$$\phi(t) = \sum_k h_k \sqrt{2}\phi(2t-k)$$

- **Does it hold for the Haar scaling function?**

$$\phi(t) = \begin{cases} \frac{1}{\sqrt{2}} & 0 \le t < 1 \\ 0 & otherwise \end{cases} \qquad h_0 = h_1 = \frac{1}{\sqrt{2}}, \quad h_k = 0 \quad for\ k > 1$$





Wavelets / Scaling function, examples – (1)

- **Given rules:** $\sum_k h_k = \sqrt{2}$ **and** $\sum_k h_k^2 = 1$

  (normalisation & orthogonality)
- **Suppose** k **=2,** $h_0 + h_1 = \sqrt{2}$

  $h_0^2 + h_1^2 = 1$
- **Results in the Haar scaling function**

$$h_0 = h_1 = \frac{1}{\sqrt{2}}$$

Wavelets / Scaling function, examples – (3)

* **Given rules:** $\sum_k h_k = \sqrt{2}$, $\sum_k h_k^2 = 1$ **and** $\sum_k h_k h_{k-2m} = \delta_m$

  (normalisation & orthogonality, like for perfect reconstruction)

* **Suppose** k **=4,** $h_0 + h_1 + h_2 + h_3 = \sqrt{2}$

$$h_0^2 + h_1^2 + h_2^2 + h_3^2 = 1$$

$$h_0 h_2 + h_1 h_3 = 0$$

* **Results: more possible solutions (Daubechies 4)…**

$$h_0 = \frac{1+\sqrt{3}}{4\sqrt{2}}, h_1 = \frac{3+\sqrt{3}}{4\sqrt{2}}, h_2 = \frac{3-\sqrt{3}}{4\sqrt{2}}, h_3 = \frac{1-\sqrt{3}}{4\sqrt{2}}$$

Wavelets / Scaling vs. Wavelet function – (1)

* **The wavelet function is orthogonal to the scaling function** $\int \phi(t-k)\psi(t-m)dt = 0$

then $w_k = \pm(-1)^k h_{N-k}$

and $\sum_k h_k w_{n-2k} = 0$

Furthermore $\sum_k w_k = 0$ **(it is high-pass!)**

Wavelets / Scaling vs. Wavelet function – (2)

* **The scaling function acts**

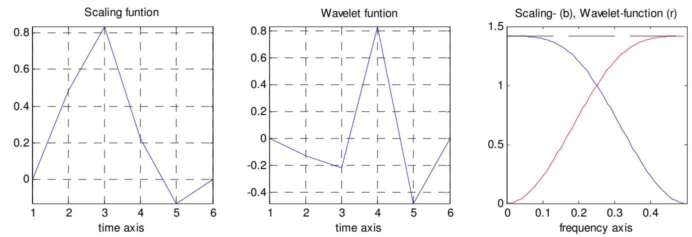
  as a low-pass (LP) filter

* **The wavelet function acts**

  as a high-pass (HP) filter

Wavelets / Daubechies 4 – (1)

Daubechy found the scaling and wavelet filters for k = 4

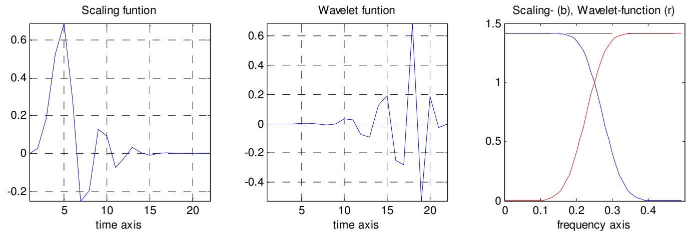


Scaling function | Wavelet function | Joint freq. response

Wavelets / Daubechies 20 – (1)

25

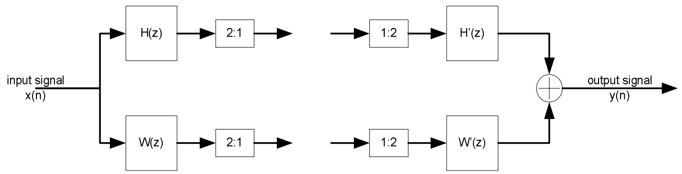
Daubechy also found the scaling and wavelet filters for k = 20



Scaling function | Wavelet function | Joint freq. response

Wavelets and filterbanks

26



Analysis (encoder):

– H, the **scaling** function

– W, the **wavelet** function

$$w_k = (-1)^k h_{n-k-1}$$

Synthesis (decoder):

– H' for LP reconstruction

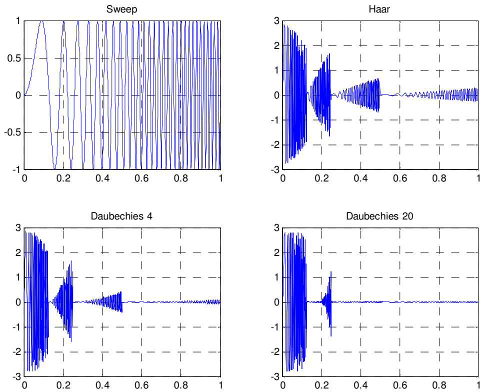
– W' for HP reconstruction

$$h_k^{-1} = \begin{cases} h_k & k = odd \\ h_{n-k-1} & k = even \end{cases}$$

$$w_k^{-1} = (-1)^k h_{n-k-1}^{-1}$$

Wavelets/ towards compression compare the energy compaction…!

27

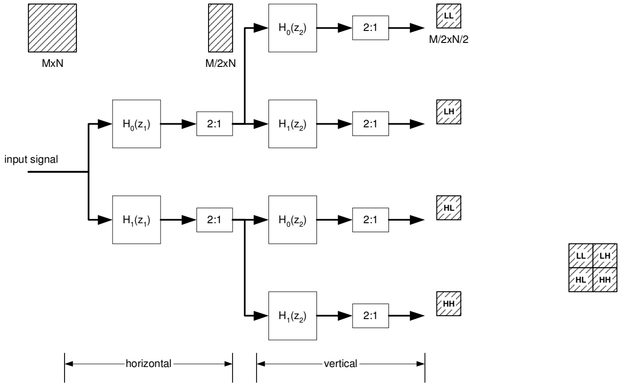


28

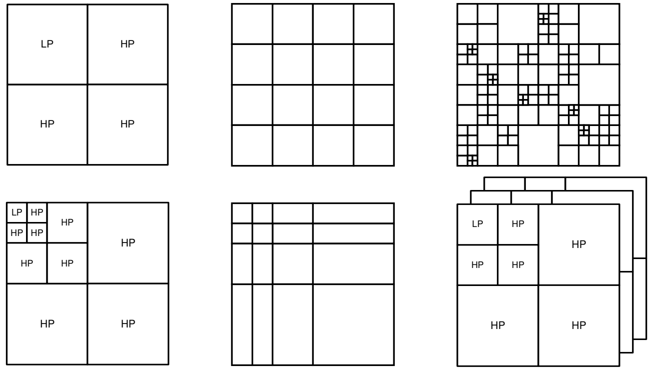
5LSE0 - Mod 04
Part 02

Introduction to Coding of Wavelet coefficients

Wavelet intraframe video coding – (1)
Two-dimensional filterbank



Wavelet intraframe video coding – (2)
Example subband decompositions

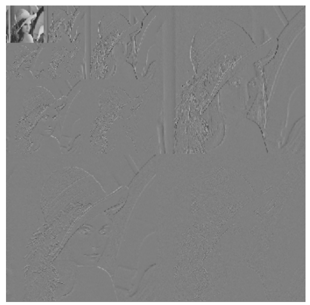


Wavelet intraframe video coding - (3)
LENA example

Wavelet intraframe video coding – (4)
LENA, 10-band decomposition

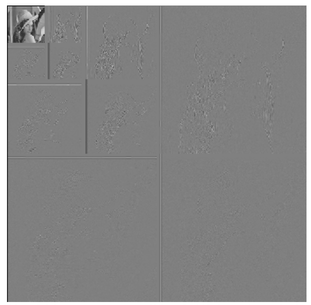
Haar wavelets

Haar transform



Daubechies-20 wavelets

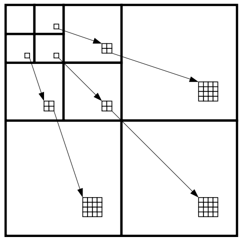
Daubechies20 transform

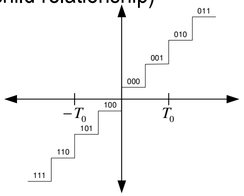


Wavelet intraframe video coding – (5)

33

* **Observation: the Wavelet (HP) sections are**
  - "relative" empty
  - related in the kind of information they carry (parent-child relationship)





Coding example, with a *N*-bit midrise quantizer: code a (complete or sub-) tree with less bits if all elements are below T_0.

Wavelet coefficient coding – (6)

Set Partitioning In Hierarchical Trees (SPIHT)

34

* **Example**

26	6	13	10
-7	7	6	4
4	-4	4	-3
2	-2	-2	0

* **Initialization:**
  - LIP = {(0,0)26, (0,1)6, (1,0)-7, (1,1)7}
  - LIS = {(0,1)D, (1,0)D, (1,1)D}
  - LSP= {}

Wavelet coefficient coding – (7)

Set Partitioning In Hierarchical Trees (SPIHT)

35

* **Initialization**
  - LIP = {(0,0)26, (0,1)6, (1,0)-7, (1,1)7}
  - LIS = {(0,1)D, (1,0)D, (1,1)D}
  - LSP= {}

26	6	13	10
-7	7	6	4
4	-4	4	-3
2	-2	-2	0

* **First pass, n=4, T=16:**
  - LIP = {(0,1)6, (1,0)-7, (1,1)7}
  - LIS = {(0,1)D, (1,0)D, (1,1)D}
  - LSP= {(0,0)26}
    - Encoding 10000 000

Wavelet coefficient coding – (8)

Set Partitioning In Hierarchical Trees (SPIHT)

36

* **First pass, n=4, T=16:**
  - LIP = {(0,1)6, (1,0)-7, (1,1)7}
  - LIS = {(0,1)D, (1,0)D, (1,1)D}
  - LSP= {(0,0)26}

26	6	13	10
-7	7	6	4
4	-4	4	-3
2	-2	-2	0

* **Second pass, n=3, T=8:**
  - LIP = {(0,1)6, (1,0)-7, (1,1)7, (1,2)6, (1,3)4}
  - LIS = {(1,0)D,(1,1)D}
  - LSP= {(0,0)26, (0,2)13, (0,3)10}
    - Encoding 000 110100000 0

Wavelet coefficient coding – (9)

Set Partitioning In Hierarchical Trees (SPIHT)

26	6	13	10
-7	7	6	4
4	-4	4	-3
2	-2	-2	0

- **Second pass, n=3, T=8:**
  - LIP = {(0,1)6, (1,0)-7, (1,1)7, (1,2)6, (1,3)4}
  - LIS = {(1,0)D,(1,1)D}
  - LSP= {(0,0)26, (0,2)13, (0,3)10}
- **Third pass, n=2, T=4:**
  - LIP = {(3,0)2, (3,1)-2, (2,3)-3, (3,2)-2, (3,3)0}
  - LIS = {}
  - LSP= {(0,0)26, (0,2)13, (0,3)10, 6, -7, 7, 6, 4, 4, -4, 4}
    - Encoding 1011101010 1101100110000 010 etc.

5LSE0 - Mod 04
Part 03

Brief introduction to JPEG2000

JPEG2000 Requirements

- **New still image coding standard for**
  - Bi-level images, like faxes and masks
  - Gray-level images, like photography and newspapers
  - Color images, like consumer & professional photography
  - Multi-component images, like in professional printing
  - Hyper-component images, like in art and prof. applications
- **Characteristics/Applications**
  - natural, medical, scientific, remote sensing, professional
- **Imaging models**
  - Client-server models, Real-time transmission
  - Image library retrieval, Imaging limited resources (embedded, mobile)

JPEG2000 Features in Part I

- **High compression efficiency** (superior to JPEG)
- **Lossless colour transformations** (continuous tone supp.)
- **Lossy and lossless coding in one algorithm**
- **Progressive by resolution, quality, position, …**
- **Static and dynamic Region-of-Interest coding/decoding**
- **Error resilience** (enable random access, robustness)
- **Perceptual quality coding** (facilitate regions of interest)
- **Multiple component image coding**
- **Tiling** (supports real-time coding, seq. build-up, limited space)
- **Light file format (optional)** (pursue open architecture, formats)

J2K Application requirements – (1)

41

- Application profiles
  - **Internet**: image 32x32 up to 4Kx4K, with 3 (Y Cr Cb, RGB, …) comp's, or 4 comp's with 1-8 bit alpha channel
  - **Printing**: compound images with typically 4800x6600 pixels (600 ppi, 8"x11") with 1, 3 or 4 components
  - **Scanning**: compound images, typically 10Kx10K up to 20Kx20K or more, with 1, 3, or 4 components and 16-bit/comp.
  - **Digital Photography**: natural, sizes up to 4K x 4K, with 1, or 3 components, and between 8-16 bits/comp.
  - **Remote sensing**: infra-red, electro-optical, multi-spectral, SAR images, horiz. res. up to 24K pixels, 1-500 components, 8-20 b/comp.
  - **Mobile**: compound, 32x32 till 4Kx4K pixels, with 1 or 3 components and 1-8 bits/component

JPEG2000

J2K Application requirements – (2)

42

- **Uncompressed**: image is stored in bit stream without compression
- **Lossless compression**: reconstructed image is identical, bit for bit, to the original, with a compression at least as good as JPEG-LS
- **Visually lossless compression**: reconstructed image differs numerically from original, the differences are not perceptible
- **Visually lossy compression**: differences are perceptible
- **Progressive Spatial Resolution**: ability to extract lower resolution images from a stream without redundant coding
- **Progressive Quality Resolution**: ability to extract lower bit-rate images without redundant coding
- **Complexity scalability**: depends on applications, different levels complexity

JPEG2000

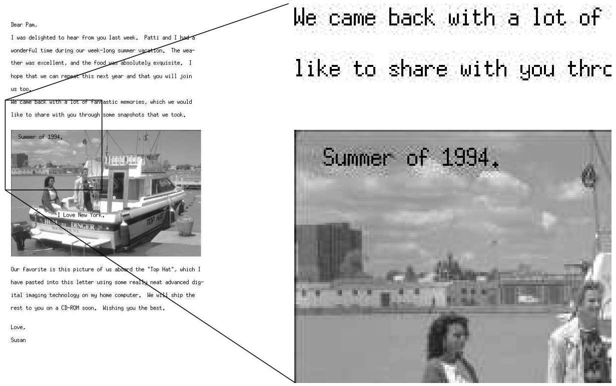
JPEG at 0.5 bpp

43

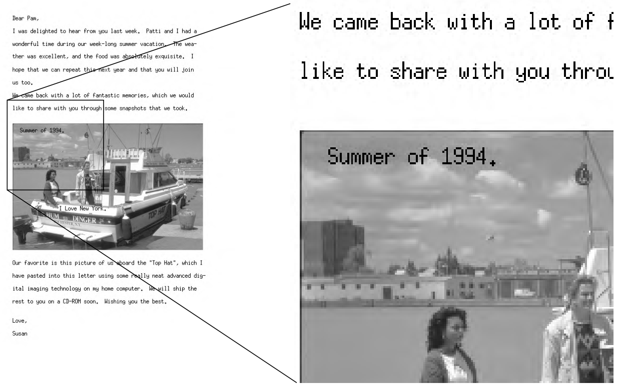
JPEG2000 at 0.5 bpp

44

JPEG compound image 1.0 bpp



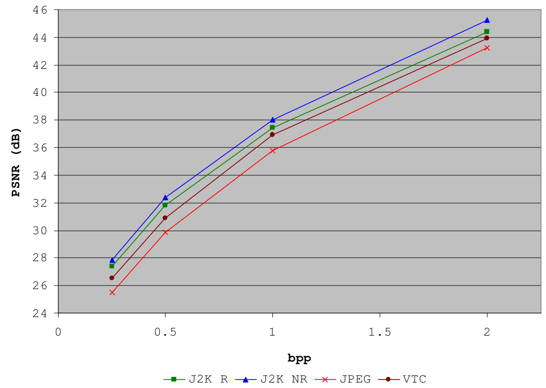
JPEG2000 compound im'g 1.0 bpp



TU/e EINDHOVEN UNIVERSITY OF TECHNOLOGY
PdW-YSu / 2022-23
Fac EE / SPS-VCA
5LSE0 / Module 04
Wavelets, DWT and main appl.
VCA

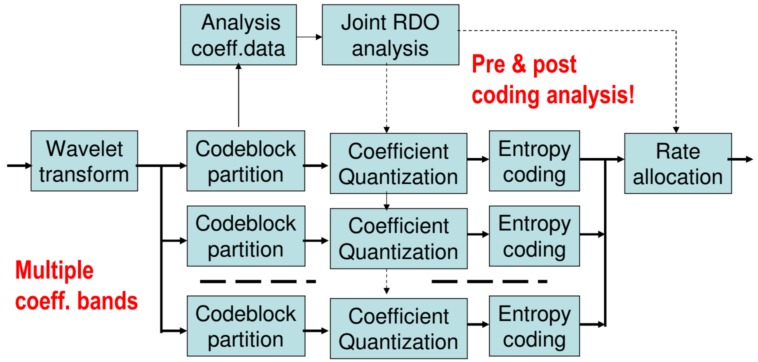
JPEG2000 and other standards

Picture quality comparison reveals: it is much better than JPEG



TU/e EINDHOVEN UNIVERSITY OF TECHNOLOGY
PdW-YSu / 2022-23
Fac EE / SPS-VCA
5LSE0 / Module 04
Wavelets, DWT and main appl.
VCA

JPEG2000 / Basic encoding scheme



TU/e EINDHOVEN UNIVERSITY OF TECHNOLOGY
PdW-YSu / 2022-23
Fac EE / SPS-VCA
5LSE0 / Module 04
Wavelets, DWT and main appl.
VCA

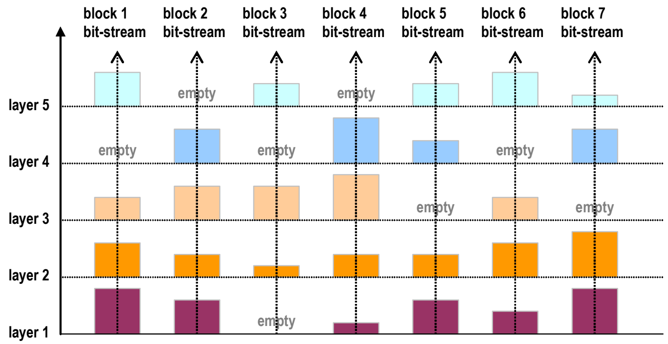
Embedded Block Coding with Optimized Truncation (EBCOT)

- Each subband is partitioned into a set of blocks (64x64 or 4x256)
- All blocks within a subband have the same size (random access)
  - (possible exception for the blocks at the image boundaries)
- Blocks are encoded independently (local variations)
- Reduce memory consumption and facilitate parallel implement.
- Post-processing operation determines the extent to which each block's bitstream should be truncated
- Final bitstream is composed of a collection of *"layers"*

EBCOT / Layered bitstream formation – (1)

- Each bitstream is organized as a succession of layers
- Each layer contains additional contributions from each block (some contributions might be empty)
- Block truncation points associated with each layer are optimal in the rate-distortion (quality-rate) sense
- Rate-distortion optimization can be done, but no standard!

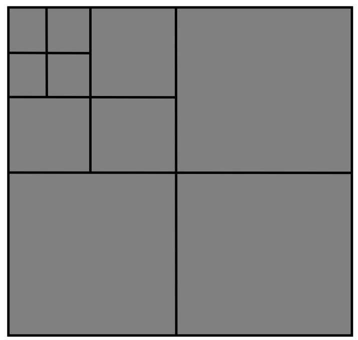
EBCOT / Layered formation – (2)



JPEG2000 / Wavelet Transform

- **Two filters supported**
  - **W9x7 (Floating point)** for lossy coding
  - **W5x3 (Integer)** for lossless coding
- **Only dyadic decomposition supported**
  - Earlier experiments also with packet and spatial decompositions
- Integer: 13x7, CRF 13x7, 5x3 etc.
- Default integer: CRF 13x7 (lossy)
- Default integer: 5x3 for lossless

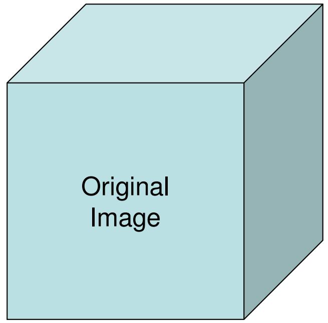
Dyadic decomposition



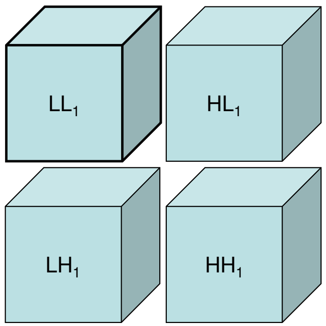
JPEG2000 Further capabilities – (1)

- **Advanced Rate Control**
  - Post-compression rate-distortion optimization
  - Requires second pass algorithm
  - Target one bit rate or an arbitrary set of specified bit rates
- **Scalability**
  - **Resolution** scalability, **amplitude** scalability, **rate** (layer) scalability
  - SNR scalability: select degree of SNR or Progr. Visual Weighting
- **Random access into bit stream**
- **Arbitrary resolutions (up to 16 levels, up to 1x1 decom)**
- **Multi-component imaging (up to 256 components)**

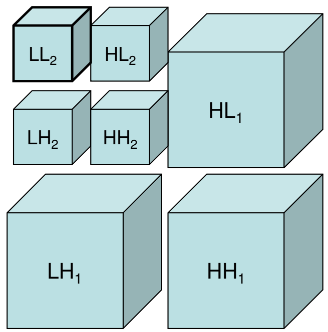
J2K Multiresolution decomposition – (1)

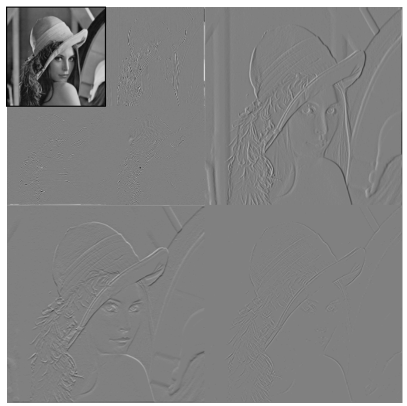


Multiresolution decomposition – (2)

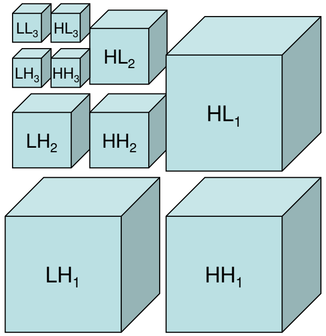


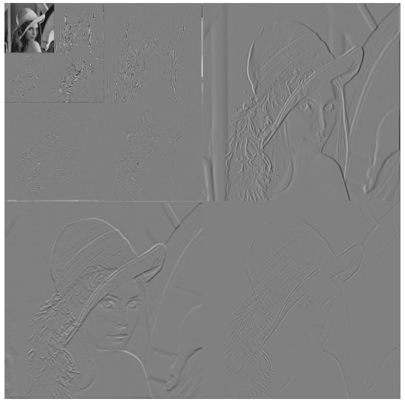
Multiresolution decomposition – (3)





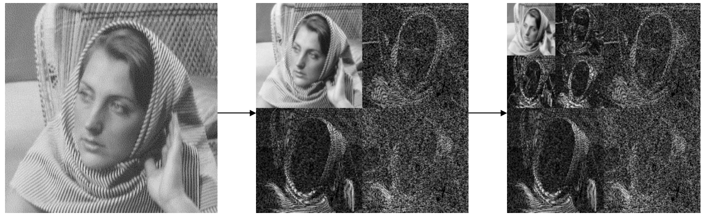
Multiresolution decomposition – (4)





Multiresolution decomposition – (5)

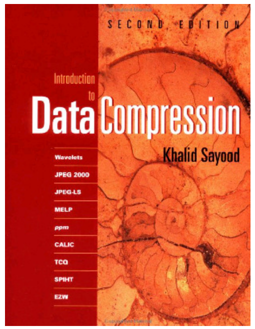
Example of dyadic decomposition into subbands

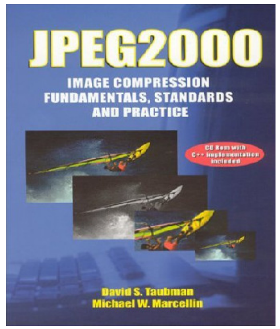


JPEG2000 / Conclusions

- **Advanced Still Image Coding System**
  - Better performance than JPEG and forms of scalability
- **Offers many functionalities, yet more complex than JPEG**
  - RDO quantization, progressive resolution & accuracy, ROI, etc.
- **No IPR associated to Part I of the standard (free licensing)**
- **Intended key standard for still image coding this decade**
  - Killer application upcoming: Digital Cinema

References & Recommended reading





...and the world wide web