Lecture 6:
Image Transformations IV:
The Discrete Wavelet Transform

Contents

Motivation

Previous Methods:

- representation in the spatial domain:
  - optimal spatial localisation
  - no direct access to frequencies or scales
- Fourier transform:
  - optimal resolution with respect to the frequencies
  - no direct access to the localisation of structures
- Laplacian pyramid is a compromise:
  - splits image into frequency bands
  - good localisation at fine scales, bad localisation at coarse scales

  However, it is redundant: requires more space than the original image.

Is there a more compact representation with localisation both in space and frequency ?

Motivation (2)

Example: Signal Representation in Another Basis

Represent the signal $\mathbf{f} = (6, 4, 5, 1)^\top$ in the following orthonormal basis of $\mathbb{R}^4$:

$$\begin{aligned}
\mathbf{b_1} &:= \frac{1}{2}(1, 1, 1, 1)^\top \\
\mathbf{b_2} &:= \frac{1}{2}(1, 1, -1, -1)^\top \\
\mathbf{b_3} &:= \frac{1}{\sqrt{2}}(1, -1, 0, 0)^\top \\
\mathbf{b_4} &:= \frac{1}{\sqrt{2}}(0, 0, 1, -1)^\top
\end{aligned}$$

Motivation (3)

In the representation $\mathbf{f} = \sum_{i=1}^{4} \alpha_i \mathbf{b_i}$ the coefficients α_i are given by

$$\begin{aligned}
\alpha_1 &= \mathbf{f}^\top \mathbf{b_1} = \frac{1}{2}(6 \cdot 1 + 4 \cdot 1 + 5 \cdot 1 + 1 \cdot 1) = 8 \\
\alpha_2 &= \mathbf{f}^\top \mathbf{b_2} = \frac{1}{2}(6 \cdot 1 + 4 \cdot 1 - 5 \cdot 1 - 1 \cdot 1) = 2 \\
\alpha_3 &= \mathbf{f}^\top \mathbf{b_3} = \frac{1}{\sqrt{2}}(6 \cdot 1 - 4 \cdot 1 + 5 \cdot 0 + 1 \cdot 0) = \frac{2}{\sqrt{2}} \\
\alpha_4 &= \mathbf{f}^\top \mathbf{b_4} = \frac{1}{\sqrt{2}}(6 \cdot 0 + 4 \cdot 0 + 5 \cdot 1 - 1 \cdot 1) = \frac{4}{\sqrt{2}}
\end{aligned}$$

These coefficients have the following interpretations:

α_1: rescaled average grey value
$|\alpha_2|$: contribution to low frequencies (without localisation)
$|\alpha_3|$: high frequency contribution in the left half of the signal
$|\alpha_4|$: high frequency contribution in the right half of the signal

The 1-D Haar Wavelet

General Idea Behind a Wavelet Basis

- A localised, wave-like function with mean 0 *(mother wavelet, Mutterwavelet)* is scaled and shifted.
- Besides these functions, an additional basis function with non-vanishing mean is required *(scaling function, Skalierungsfunktion)*.

M I A
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28

Alfréd Haar (1885–1933) was a Hungarian mathematician who studied in Göttingen. In his Ph.D. thesis that was supervised by David Hilbert he introduced the first wavelet concepts. He made a number of significant contributions to the field of analysis. Source: `http://www-history.mcs.st-andrews.ac.uk/PictDisplay/Haar.html`.

The 1-D Haar Wavelet (3)

The Continuous Haar Wavelet

- simplest wavelet (Alfréd Haar, 1910)
- uses a step function as mother wavelet:

$$\Psi(x) := \begin{cases} 1 & \text{for } 0 \le x \le \frac{1}{2}, \\ -1 & \text{for } \frac{1}{2} < x \le 1, \\ 0 & \text{else} \end{cases}$$

- consider scaled and shifted versions:

$$\Psi_{j,k}(x) \;:=\; \frac{1}{2^{j/2}}\,\Psi\left(\frac{x}{2^j} - k\right).$$

$\Psi_{j,k}$ has width 2^j, height $\frac{1}{2^{j/2}}$, and starts at $k2^j$.
The scale level is specified by j, the shift by k.
Finer scales correspond to smaller scale levels j.

The 1-D Haar Wavelet (4)

- The factor $\frac{1}{2^{j/2}}$ guarantees that $\Psi_{j,k}$ has norm 1,

$$\|\Psi_{j,k}\| \;:=\; \sqrt{\langle \Psi_{j,k}, \Psi_{j,k}\rangle} \;=\; 1,$$

in the space of quadratically (Lebesgue) integrable functions

$$L^2(\mathbb{R}) \;:=\; \left\{ f:\mathbb{R}\to\mathbb{R} \;\middle|\; \int_{-\infty}^{\infty} |f(x)|^2\,dx < \infty \right\}$$

with the inner product

$$\langle f, g\rangle \;:=\; \int_{-\infty}^{\infty} f(x)\,g(x)\,dx$$

- If j and k are integer numbers, the Haar wavelets $\{\Psi_{j,k}\}$ are even orthonormal:

$$\langle \Psi_{j,k}\,, \Psi_{n,m}\rangle \;=\; \begin{cases} 1 & \text{for } (j,k) = (n,m), \\ 0 & \text{else.} \end{cases}$$

The 1-D Haar Wavelet (5)

- As scaling function one chooses a box function:

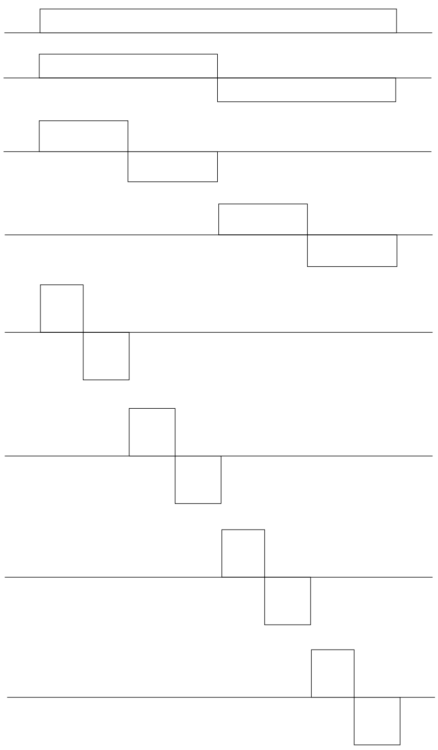
$$\Phi(x) := \begin{cases} 1 & \text{for } 0 \le x \le 1, \\ 0 & \text{else.} \end{cases}$$

- It may also be scaled and shifted:

$$\Phi_{j,k}(x) := \frac{1}{2^{j/2}} \Phi\left(\frac{x}{2^j} - k\right).$$

Also $\Phi_{j,k}$ has width 2^j, height $\frac{1}{2^{j/2}}$, and starts at $k2^j$.

The 1-D Haar Wavelet (6)



Scaling function and Haar wavelets for $n = 3$. We consider the interval $[0, 2^3] = [0, 8]$. **From top to bottom:** Scaling function $\Phi_{3,0}$, wavelets $\Psi_{3,0}$, $\Psi_{2,0}$, $\Psi_{2,1}$, $\Psi_{1,0}$, $\Psi_{1,1}$, $\Psi_{1,2}$, $\Psi_{1,3}$. These 8 orthonormal functions allow to represent every discrete signal of length 8. Author: S. Zimmer (2002).

MIA 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28

The 1-D Haar Wavelet (7)

The Discrete Haar Wavelet

- The previous model was continuous in order to describe scaling and shifting in a more intuitive way. Discrete signals require discrete models.

- For a discrete signal of length $N = 2^n$ one considers the N functions

$\Phi_{n,0}$,	scaling function
$\Psi_{n,0}$,	lowest frequency, unlocalised
$\Psi_{n-1,0}$, $\Psi_{n-1,1}$	higher frequency, at 2 locations
$\vdots \quad \ddots$	
$\Psi_{1,0}$, $\Psi_{1,1}$, ... ,$\Psi_{1,2^{n-1}-1}$	highest frequency, at 2^{n-1} locations.

  Sampling at N equidistant grid points $\{\frac{1}{2}, \frac{3}{2}, ..., N-\frac{1}{2}\}$ creates an orthonormal basis of $\mathbb{R}^N$ (see assignment).

- We identify the step functions $\Psi_{j,k}(x)$ and $\Phi_{j,k}(x)$ with their vectors $\mathbf{\Psi_{j,k}}$ and $\mathbf{\Phi_{j,k}}$.

The 1-D Haar Wavelet (8)

M I A
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28

Example

- In the example on Page 3 we had $n = 2$, yielding the basis vectors

$$\mathbf{\Phi_{2,0}} = \tfrac{1}{2}\,(1,1,1,1)^\top,$$
$$\mathbf{\Psi_{2,0}} = \tfrac{1}{2}\,(1,1,-1,-1)^\top,$$
$$\mathbf{\Psi_{1,0}} = \tfrac{1}{\sqrt{2}}\,(1,-1,0,0)^\top, \qquad \mathbf{\Psi_{1,1}} = \tfrac{1}{\sqrt{2}}\,(0,0,1,-1)^\top.$$

Widely Used Conventions

- $c_{j,k}$: coefficient of the scaling vector $\mathbf{\Phi_{j,k}}$ (c like *coarse*)

- $d_{j,k}$: coefficient for the wavelet vector $\mathbf{\Psi_{j,k}}$ (d like *detail*)

- coefficients are stored in a coarse-to-fine manner:
  $c_{n,0} \;|\; d_{n,0} \;|\; d_{n-1,0} \;\; d_{n-1,1} \;|\; \dots \;|\; d_{1,0} \;\; \dots \;\; d_{1,2^{n-1}-1}$

The Fast Wavelet Transform

Important Practical Aspect

- Naive implementation of the discrete wavelet transform requires $\mathcal{O}(N^2)$ operations for a signal of length N.
  Are there algorithms with lower complexity?

M I A
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28

The Fast Wavelet Transform (FWT)

- The definitions of the Haar wavelets and the scaling functions imply

$$\mathbf{\Phi_{j,k}} = \frac{1}{\sqrt{2}}(\mathbf{\Phi_{j-1,2k}} + \mathbf{\Phi_{j-1,2k+1}})$$

$$\mathbf{\Psi_{j,k}} = \frac{1}{\sqrt{2}}(\mathbf{\Phi_{j-1,2k}} - \mathbf{\Phi_{j-1,2k+1}})$$

  as well as

$$f_k = \mathbf{f}^\top \mathbf{\Phi_{0,k}} \qquad (k = 0, ..., N-1).$$

- Because of

$$c_{j,k} = \mathbf{f}^\top \mathbf{\Phi_{j,k}},$$
$$d_{j,k} = \mathbf{f}^\top \mathbf{\Psi_{j,k}},$$

  these relations for the basis vectors carry over to the coefficients.

The Fast Wavelet Transform (3)

- Thus, for the scales $j = 1, ..., n$ we compute

$$
\begin{aligned}
c_{j,k} &= \frac{1}{\sqrt{2}} (c_{j-1,2k} + c_{j-1,2k+1}) \\
d_{j,k} &= \frac{1}{\sqrt{2}} (c_{j-1,2k} - c_{j-1,2k+1})
\end{aligned}
$$

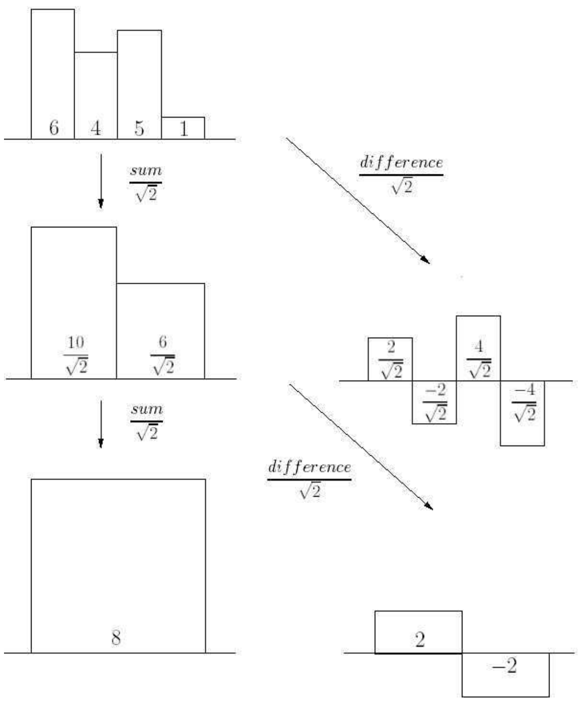
  starting from the finest scale with the initialisation

$$c_{0,k} = f_k \qquad (k = 0, ..., N-1).$$

- This fine-to-coarse algorithm is called *Fast Wavelet Transform (FWT)*.
  It resembles the Laplacian pyramid decomposition from Lecture 5.
  Its complexity is $\mathcal{O}(N)$.

The Fast Wavelet Transform (4)

M I A
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28



Pyramid-like representation of the FWT of the signal $(6, 4, 5, 1)^\top$. The left part resembles the Gaussian pyramid and gives the scaling coefficient $c_{2,0} = 8$. The right part resembles the Laplacian pyramid and yields the wavelet coefficients $d_{1,0} = \frac{2}{\sqrt{2}}$, $d_{1,1} = \frac{4}{\sqrt{2}}$, $d_{2,0} = 2$. Author: S. Zimmer (2002).

The Fast Wavelet Transform (5)

The Inverse Fast Wavelet Transformation

- Because of

$$\mathbf{\Phi_{j,2k}} = \frac{1}{\sqrt{2}}(\mathbf{\Phi_{j+1,k}} + \mathbf{\Psi_{j+1,k}})$$

$$\mathbf{\Phi_{j,2k+1}} = \frac{1}{\sqrt{2}}(\mathbf{\Phi_{j+1,k}} - \mathbf{\Psi_{j+1,k}})$$

  the inverse transformation is as simple as the forward transformation:

- Proceed in a coarse-to-fine manner from $j = n-1$ to $j = 0$ and compute

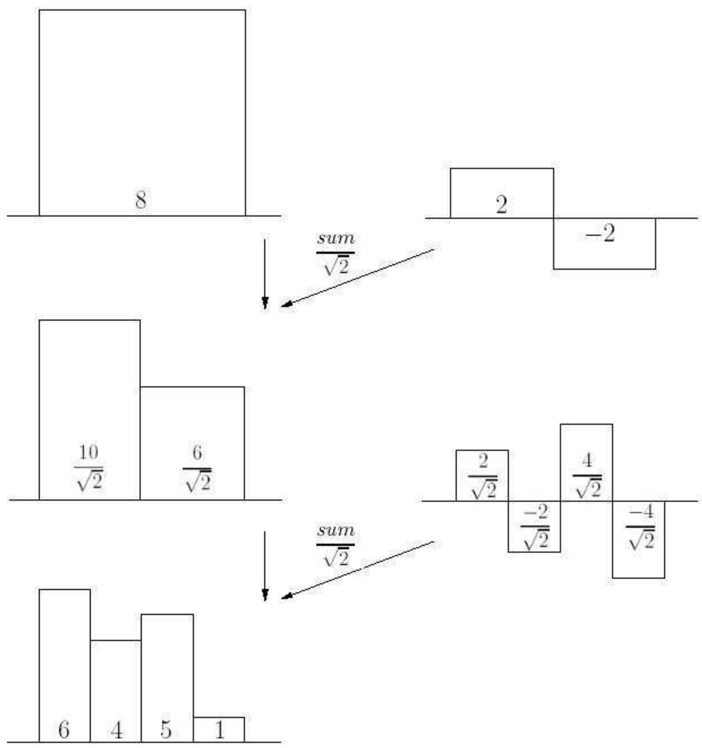
$$c_{j,2k} = \frac{1}{\sqrt{2}}(c_{j+1,k} + d_{j+1,k})$$

$$c_{j,2k+1} = \frac{1}{\sqrt{2}}(c_{j+1,k} - d_{j+1,k})$$

- Then the reconstructed signal $\mathbf{f}$ is given by $f_k = c_{0,k}$.

MI A
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28

The Fast Wavelet Transform (6)



Reconstruction of the original signal $(6, 4, 5, 1)^\top$ starting from the scaling coefficient $c_{2,0} = 8$ and the wavelet coefficients $d_{2,0} = 2$, $d_{1,0} = \frac{2}{\sqrt{2}}$, $d_{1,1} = \frac{4}{\sqrt{2}}$. Author: S. Zimmer (2002).

Wavelets Versus Pyramids and Fourier Representations

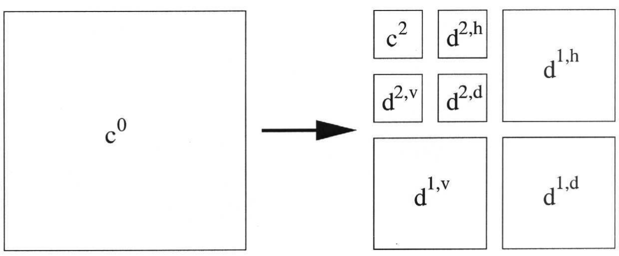
- The discrete wavelet coefficients can be computed in optimal complexity: $\mathcal{O}(N)$. (in contrast to FFT: $\mathcal{O}(N \log_2 N)$)
- Pyramids and the discrete wavelet transform are not shift invariant!
- Unlike the pyramid decomposition, the discrete wavelet representation has no redundancy: A signal of length N is represented by N coefficients.

M I A
1 2
3 4
5 6
7 8
9 10
11 12
13 14
15 16
17 18
19 20
21 22
23 24
25 26
27 28

The Two-Dimensional Wavelet Transform

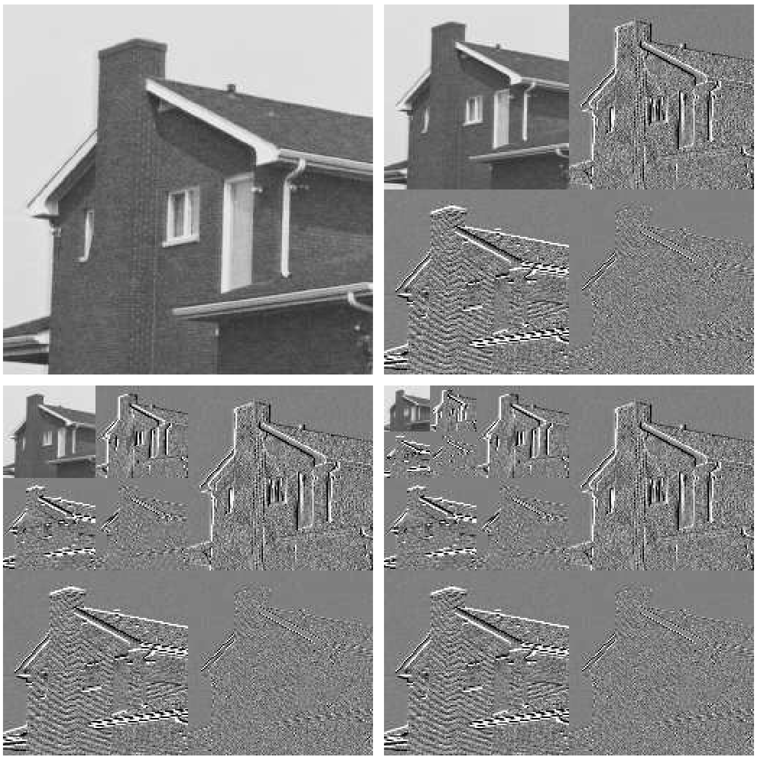
Frequently Used (Nonstandard Decomposition)

- Start with computing the wavelet decomposition on a *single* level, first in x direction then in y direction.
- Perform the next decomposition only in the quadrant that contains the low-frequent parts (scaling coefficients) from both directions.
- Proceed until a single pixel is reached.



Distribution of the coefficients after two decomposition steps. Author: C. Schnörr (1999).

Two-Dimensional Wavelet Transform (2)



Two-dimensional nonstandard wavelet decomposition. **(a) Top left:** Original image, 256×256 pixels. **(b) Top right:** After 1 decomposition step. **(c) Bottom left:** After 2 decomposition steps. **(d) Bottom right:** After 3 decomposition steps. Author: J. Weickert (2002).

Application to Data Compression (1)

Application to Data Compression

- In images, one can cancel wavelet coefficients that are small in magnitude without introducing severe visual degradations.
- Because of their locality, wavelets belong to the best bases for this purpose. Therefore they are used in modern compression standards such as JPEG 2000.

M I A
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28

Application to Data Compression (2)

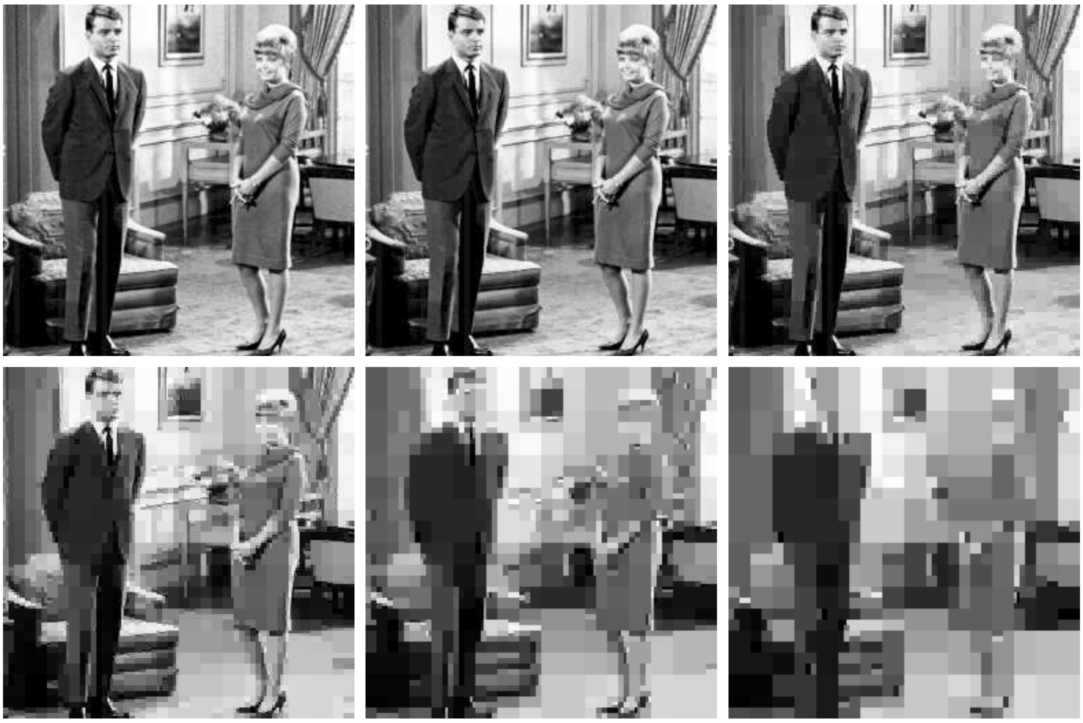


Image compression by removing the Haar wavelet coefficients that are smallest in magnitude. **(a) Top left:** Original image, 256×256 pixels. **(b) Top middle:** 66.37 % of all coefficients removed. **(c) Top right:** 90.34 % **(d) Bottom left:** 96.68 %. **(e) Bottom middle:** 99.01 %. **(f) Bottom right:** 99.66 %. Author: J. Weickert (2002).

Summary (1)

Summary

- Wavelets provide a signal representation that is localised in space and frequency.
- simplest wavelet: Haar wavelet
- The fast wavelet transform (FWT) is similar to the Laplacian pyramid and has linear complexity.
- In higher dimensions one often uses the so-called nonstandard decomposition.
- most important wavelet application: data compression
- In general, wavelets are neither shift invariant nor invariant under rotations.

M I A
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28

Literature

- R. C. Gonzalez, R. E. Woods: *Digital Image Processing*. Prentice Hall, Upper Saddle River, Second Edition, 2002.
  (Chapter 7 deals with wavelets.)
- E. J. Stollnitz, T. D. DeRose, D. H. Salesin: *Wavelets for Computer Graphics*. Morgan Kaufmann, San Francisco, 1996.
  (contains a well-readable introduction to Haar wavelets)
- W. Bäni: *Wavelets*. Oldenbourg, München, 2002.
  (fairly simple introduction to wavelet concepts; in German)
- S. Mallat: *A Wavelet Tour of Signal Processing*. Academic Press, San Diego, Second Edition, 1999.
  (wavelet bible)
- A. Haar: Zur Theorie der orthogonalen Funktionensysteme. *Mathematische Annalen*, Vol. 69, pp. 331–371, 1910.
  (intriduced the Haar wavelet)

Assignment C2 – Classroom Work

Problem 1 (Discrete Wavelet Transform)

Let $n \in \mathbb{N}$ and $N = 2^n$. Show that the vectors

$$\begin{array}{l} \Phi_{n,0}, \\ \Psi_{n,0}, \\ \Psi_{n-1,0},\ \Psi_{n-1,1}, \\ \vdots \qquad\qquad \ddots \\ \Psi_{1,0},\ \Psi_{1,1},\ \dots,\Psi_{1,2^{n-1}-1} \end{array}$$

as defined in this lecture form an orthonormal basis of $\mathbb{R}^N$ with respect to the Euclidean inner product, i.e. show that

- they are N vectors.
- different vectors are orthogonal.
- they have norm 1.

(*This property allows to interpret the wavelet transform as a change of basis.*)

M I A
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28

Assignment T2 – Theoretical Homework

Problem 1 (Image Pyramids) (6 points)

Calculate the Gaussian and the Laplacian pyramid of the signal

$$f := (-20, -2, 11, -4, -12, -20, 6, 24, 12)$$

Reconstruct the initial signal from the Laplacian pyramid.

(*This problem will be useful for understanding the idea behind image pyramids.*)

Problem 2 (Discrete Fourier Transform) (8 points)

Let the following signal f_1 and its shifted variant f_2 be given:

$$f_1 := (6, 8, 2, 4)\ , \qquad f_2 := (4, 6, 8, 2)\ .$$

(a) Compute the DFT and the spectra of both signals and compare them. What are your findings ?

(b) Set the highest frequency in the DFT of f_1 to zero and transform the result back. What can you observe with respect to the smoothness of the signal ?

(*This problem will be useful for understanding some properties of the DFT.*)

Problem 3 (Discrete Fourier Transform) (6 points)

For complex vectors $f = (f_i)_{i=0}^{M-1}$ and $g = (g_i)_{i=0}^{M-1}$, one defines their Hermitian inner product as $\langle f, g \rangle := \sum_{m=0}^{M-1} f_m \bar{g}_m$ where $\bar{g}_m$ is the complex conjugate of g_m.

Show that with respect to this inner product, the M vectors

$$v_p := \frac{1}{\sqrt{M}} \left(\exp\left(\frac{2\pi i\, p\, 0}{M}\right), \exp\left(\frac{2\pi i\, p\, 1}{M}\right), \ldots, \exp\left(\frac{2\pi i\, p\, (M-1)}{M}\right) \right)^{\top}$$

with $p = 0,\ldots,M-1$ form an orthonormal basis of the M-dimensional complex vector space $\mathbb{C}^M$.

(*This property allows to interpret the DFT as a change of basis.*)

Deadline for submission: Tuesday, November 20, 10 am (before the lecture).