Wavelets and Multiresolution Processing
(Wavelet Transforms)

Christophoros Nikou

cknikou@cs.uoi.gr

Contents

– Image pyramids
– Subband coding
– The Haar transform
– Multiresolution analysis
  • Series expansion
  • Scaling functions
  • Wavelet functions
– Wavelet series
– Discrete wavelet transform (DWT)
– Fast wavelet transform (FWT)
– Wavelet packets
1-D Wavelet Transforms

The Wavelet Series

A continuous signal may be represented by a scaling function in a subspace $V_{j_0}$ and some number of wavelet functions in subspaces $W_{j_0}, W_{j_0+1}, W_{j_0+2}, \ldots$

\[
f(x) = \sum_{k} c_{j_0}(k) \phi_{j_0,k}(x) + \sum_{j=j_0}^{\infty} \sum_{k} d_{j}(k) \psi_{j,k}(x)
\]

Scaling coefficients
\[
c_{j_0}(k) = \left\langle f(x), \phi_{j_0,k}(x) \right\rangle = \int f(x) \phi_{j_0,k}(x) \, dx
\]

Detail (wavelet) coefficients
\[
d_{j}(k) = \left\langle f(x), \psi_{j,k}(x) \right\rangle = \int f(x) \psi_{j,k}(x) \, dx
\]

Example: using Haar wavelets and starting from $j_0=0$, compute the wavelet series of

\[
f(x) = \begin{cases} x^2, & 0 \leq x \leq 1 \\ 0, & \text{otherwise} \end{cases}
\]

Scaling coefficients
\[
c_0(0) = \int_0^1 x^2 \phi_{0,0}(x) \, dx = \frac{1}{3}
\]

There is only one scaling coefficient for $k=0$. Integer translations of the scaling function do not overlap with the signal.
Example (continued): Detail (wavelet) coefficients

\[ d_0(0) = \int_{0}^{1} x^2 \psi_{0,0}(x) \, dx = \int_{0}^{0.5} x^2 \, dx - \int_{0.5}^{1} x^2 \, dx = -\frac{1}{4} \]

\[ d_1(0) = \int_{0}^{1} x^2 \psi_{1,0}(x) \, dx = \int_{0}^{0.25} x^2 \sqrt{2} \, dx - \int_{0.25}^{0.5} x^2 \sqrt{2} \, dx = -\frac{\sqrt{2}}{32} \]

\[ d_1(1) = \int_{0}^{1} x^2 \psi_{1,1}(x) \, dx = \int_{0}^{0.75} x^2 \sqrt{2} \, dx - \int_{0.75}^{1} x^2 \sqrt{2} \, dx = -\frac{3\sqrt{2}}{32} \]

Example (continued): Substituting these values:

\[ f(x) = \frac{1}{3} \phi_{0,0}(x) + \left[ -\frac{1}{4} \psi_{0,0}(x) \right] + \left[ -\frac{\sqrt{2}}{32} \psi_{1,0}(x) - \frac{\sqrt{2}}{32} \psi_{1,1}(x) \right] + \ldots \]

\[ V_0 = V_0 \oplus W_0 \]
\[ V_1 = V_0 \oplus W_0 \oplus W_1 \]
\[ V_2 = V_1 \oplus W_1 \oplus W_2 \oplus W_1 \]

C. Nikou – Digital Image Processing (E12)
### Example (continued):

The scaling function approximates the signal by its average value.

Each wavelet subspace adds a level of detail in the wavelet series representation of the signal.

---

### 1-D Wavelet Transforms

#### The Discrete Wavelet Transform

If the signal is discrete, of length $M$, we compute its DWT.

$$f(n) = \frac{1}{\sqrt{M}} \sum_k W_\phi(j_0, k) \phi_{j_0, k}(n) + \frac{1}{\sqrt{M}} \sum_{j=j_0}^{\infty} \sum_k W_\psi(j, k) \psi_{j, k}(n)$$

**Scaling coefficients**

$$W_\phi(j_0, k) = \frac{1}{\sqrt{M}} \sum_{n=0}^{M-1} f(n) \phi_{j_0, k}(n)$$

**Detail coefficients**

$$W_\psi(j, k) = \frac{1}{\sqrt{M}} \sum_{n=0}^{M-1} f(n) \psi_{j, k}(n), \text{ for } j \geq j_0$$
The Discrete Wavelet Transform (cont.)

\[ f(n) = \frac{1}{\sqrt{M}} \sum_{k} W_\phi(j_0, k) \phi_{j_0, k}(n) + \frac{1}{\sqrt{M}} \sum_{j=j_0}^{\infty} \sum_{k} W_\psi(j, k) \psi_{j, k}(n) \]

We take \( M \) equally spaced samples over the support of the scaling and wavelet functions. Normally, we let \( j_0=0 \) and \( M=2^J \).

Therefore, the summations are performed over \( n=0, 1, 2, \ldots, M-1 \), \( j=0, 1, 2, \ldots, J-1 \), \( k=0, 1, 2, \ldots, 2^j-1 \)

For Haar wavelets, the discretized scaling and wavelet functions correspond to the rows of the \( M \times M \) Haar matrix.

---

**Example:** Let \( f(n) = \{1, 4, -3, 0\} \).

We will compute the DWT of the signal using the Haar scaling function and the corresponding wavelet functions.

Here, \( M=4=2^J \), \( J=2 \) and with \( j_0=0 \) the summations are performed over \( n=0, 1, 2, 3 \), \( j=0, 1 \) and \( k=0 \) for \( j=0 \) or \( k=0, 1 \) for \( j=1 \).

The values of the sampled scaling and wavelet functions are the elements of the rows of \( H_4 \).
Example (continued): \( f(n) = \{1, 4, -3, 0\} \).

The DWT relative to the Haar scaling and wavelet functions is

\[
\begin{align*}
W_\phi(j_0, k) &= \frac{1}{\sqrt{M}} \sum_{n=0}^{M-1} f(n) \phi_{j_0,k}(n) = \frac{1}{2} \sum_{n=0}^{3} f(n) \phi_{0,0}(n) \\
W_\phi(0, 0) &= \frac{1}{2} [1 \cdot 1 + 4 \cdot 1 - 3 \cdot 1 + 0 \cdot 1] = 1 \\
W_\psi(0, 0) &= \frac{1}{2} [1 \cdot 1 + 4 \cdot 1 - 3 \cdot (-1) + 0 \cdot (-1)] = 4 \\
W_\psi(1, 0) &= \frac{1}{2} [1 \cdot \sqrt{2} + 4 \cdot (-\sqrt{2}) - 3 \cdot 0 + 0 \cdot 0] = -1.5\sqrt{2} \\
W_\psi(1, 1) &= \frac{1}{2} [1 \cdot 0 + 4 \cdot 0 - 3 \cdot \sqrt{2} + 0 \cdot (-\sqrt{2})] = -1.5\sqrt{2}
\end{align*}
\]

To reconstruct the signal, we compute

\[
f(n) = \frac{1}{\sqrt{M}} \sum_k W_\phi(j_0, k) \phi_{j_0,k}(n) + \frac{1}{\sqrt{M}} \sum_{j=j_0}^{\infty} \sum_k W_\psi(j, k) \psi_{j,k}(n) \\
= \frac{1}{2} \left[ W_\phi(0, 0) \phi_{0,0}(n) + W_\psi(0, 0) \psi_{0,0}(n) + W_\psi(1, 0) \psi_{1,0}(n) + W_\psi(1, 1) \psi_{1,1}(n) \right]
\]

for \( n = 0, 1, 2, 3 \). Notice that we could have started from a different approximation level \( j_0 \neq 0 \).
A computationally efficient implementation of the DWT [Mallat 1989]. It resembles subband coding. Consider the multiresolution refinement equation

\[ \phi(x) = \sum_n h_\phi(n)\sqrt{2}\phi(2x - n) \]

Scaling \( x \) by \( 2^j \), translating it by \( k \) and letting \( m = 2k + n \):

\[ \phi(2^j x - k) = \sum_n h_\phi(n)\sqrt{2}\phi(2(2^j x - k) - n) = \]

\[ = \sum_n h_\phi(n)\sqrt{2}\phi(2^{j+1} x - 2k - n) \]

\[ = \sum_m h_\phi(m - 2k)\sqrt{2}\phi(2^{j+1} x - m) \]

The scaling vector \( h_\phi \) may be considered as weights used to expand \( \phi(2^j x - k) \) as a sum of scale \( j+1 \) scaling functions.

A similar sequence of operations leads to

\[ \psi(2^j x - k) = \sum_m h_\psi(m - 2k)\sqrt{2}\phi(2^{j+1} x - m) \]
Injecting these expressions into the DWT formulas

\[ W_\phi(j, k) = \frac{1}{\sqrt{M}} \sum_{n=0}^{M-1} f(n) \phi_{j,k}(n) \]

\[ W_\psi(j, k) = \frac{1}{\sqrt{M}} \sum_{n=0}^{M-1} f(n) \psi_{j,k}(n), \quad \text{for } j \geq j_0 \]

yields the following important equations relating the DWT coefficients of adjacent scales.

\[ W_\phi(j, k) = \sum_m h_\phi(m-2k) W_\phi(j+1, m) \]

\[ W_\psi(j, k) = \sum_m h_\psi(m-2k) W_\phi(j+1, m) \]

Both the scaling and the wavelet coefficients of a certain scale \( j \) may be obtained by

- convolution of the scaling coefficients of the next scale \( j+1 \) (the finer detail scale), with the order-reversed scaling and wavelet vectors \( h_\phi(-n) \) and \( h_\psi(-n) \).
- subsampling the result.

The complexity is \( O(M) \).
1-D Wavelet Transforms
The Fast Wavelet Transform (cont...)

\[
W_{\phi}(j, k) = h_{\phi}(-n) * W_{\phi}(j+1, n) \big|_{n=2k, k \geq 0}
\]

\[
W_{\psi}(j, k) = h_{\psi}(-n) * W_{\phi}(j+1, n) \big|_{n=2k, k \geq 0}
\]

The convolutions are evaluated at non-negative even indices. This is equivalent to filtering and downsampling by 2.

The procedure may be iterated to create multistage structures of more scales. The highest scale coefficients are assumed to be the values of the signal itself \( W_{\phi}(J, n) = f(n) \).
The corresponding frequency splitting characteristics for the two-stage procedure.

C. Nikou – Digital Image Processing (E12)
1-D Wavelet Transforms
The Fast Wavelet Transform (cont...)

Reconstruction may be obtained by the IFWT
• upsampling by 2 (inserting zeros).
• convolution of the scaling coefficients by $h_\phi(n)$ and the wavelet coefficients by $h_\psi(n)$ and summation.

\[
W_\phi(j + 1, k) = h_\phi(k) \ast W_\phi^{\uparrow 2} (j, k) + h_\psi(k) \ast W_\psi^{\uparrow 2} (j, k)_{k \geq 0}
\]

C. Nikou – Digital Image Processing (E12)

---

1-D Wavelet Transforms
The Fast Wavelet Transform (cont...)

It can also be extended to multiple stages.
1-D Wavelet Transforms
The Fast Wavelet Transform (cont...)

- The Fourier basis functions guarantee the existence of the transform for energy signals.
- The wavelet transform depends upon the availability of scaling functions for a given wavelet function.
- The wavelet transform depends on the orthonormality (biorthogonality) of the scaling and wavelet functions.
- The F.T. informs us about the frequency content of a signal. It does not inform us on the specific time instant that a certain frequency occurs.
- The W.T. provides information on “when the frequency occurs”.
1-D Wavelet Transforms
Relation to the Fourier Transform (cont…)

Heisenberg cells.
• The width of each rectangle in (a) represents one time instant.
• The height of each rectangle in (b) represents a single frequency.

Time domain pinpoints the instance an event occurs but has no frequency information. DFT domain pinpoints the frequencies that are present in the events but provides no time resolution (when a certain frequency appears).
1-D Wavelet Transforms
Relation to the Fourier Transform (cont...)

In DWT the time-frequency resolution varies but the area of each tile is the same. There is a compromise between time and frequency resolutions.

At low frequencies the tiles are shorter (better frequency resolution, less ambiguity regarding frequency) but wider (more ambiguity regarding time).

At high frequencies the opposite happens (time resolution is improved).

2-D Discrete Wavelet Transform

Extension from 1-D wavelet transforms. A 2-D scaling function and three 2-D wavelet functions are required.

They are separable. They are the product of two 1-D functions.

\[
\phi(x, y) = \phi(x)\phi(y)
\]

Variations along columns

\[
\psi^H(x, y) = \psi(x)\phi(y)
\]

Variations along rows

\[
\psi^V(x, y) = \phi(x)\psi(y)
\]

Variations along diagonals

\[
\psi^D(x, y) = \psi(x)\psi(y)
\]
Given separable scaling and wavelet functions the scaled and translated basis

\[ \phi_{j,m,n}(x, y) = 2^{j/2} \phi(2^j x - m, 2^j y - n) \]
\[ \psi^i_{j,m,n}(x, y) = 2^{j/2} \psi^i(2^j x - m, 2^j y - n), \; i = \{H, V, D\} \]

where \( i \) is an index and there are two translations \( m \) and \( n \).

The DWT is given by

\[ W_\phi(j_0, m, n) = \frac{1}{\sqrt{MN}} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} f(m, n) \phi_{j_0,m,n}(m, n) \]
\[ W^i_{\psi}(j, m, n) = \frac{1}{\sqrt{MN}} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} f(m, n) \psi^i_{j_0,m,n}(m, n), \; i = \{H, V, D\} \]

And the inverse transformation is

\[ f(m, n) = \frac{1}{\sqrt{MN}} \sum_m \sum_n W_\phi(j_0, m, n) \phi_{j_0,m,n}(m, n) \]
\[ + \frac{1}{\sqrt{MN}} \sum_{j=1}^{J} \sum_{i=H}^{2} \sum_{k=0}^{K} \sum_n \sum_m W^i_{\psi}(j, m, n) \psi^i_{j,m,n}(m, n) \]
2-D Discrete Wavelet Transform (cont...)

We take $M$ equally spaced samples over the support of the scaling and wavelet functions. We normally, we let $j_0=0$ and select $N=M=2^J$ so that

$$m=n=0, 1, 2, ..., M-1, \quad j=0, 1, 2, ..., J-1, \quad k=0, 1, 2, ..., 2^J-1$$

Like the 1-D transform it can be implemented using filtering and downsampling. We take the 1-D FWT of the rows followed by the 1-D FWT of the resulting columns.
Reconstruction (inverse 2-D FWT)

C. Nikou – Digital Image Processing (E12)

2-D Discrete Wavelet Transform (cont…)

FIGURE 7.25
Comparing 2-D three-scale FWT:
(a) the original image (b) one-scale FWT (c) a
pre-scale FWT and (d) a three-scale FWT.

C. Nikou – Digital Image Processing (E12)
2-D Discrete Wavelet Transform (cont…)

a) Noisy CT image. We wish to remove the noise by manipulating the DWT.

c) Zeroing the highest resolution detail coefficients (not thresholding the lowest resolution details).

e) Zeroing all of the detail coefficients of the two scale DWT.

b) Thresholding all of the detail coefficients of the two scale DWT. Significant edge information is eliminated.

d) Information loss in c) with respect to the original image. The edges are slightly disturbed.

d) Information loss in e) with respect to the original image. Significant edge information is removed.
Wavelet Packets

The DWT, as it was defined, decomposes a signal into a sum of scaling and wavelet functions whose bandwidths are logarithmically related.

• low frequencies use narrow bandwidths.
• high frequencies use wider bandwidths.

For greater flexibility of the partitioning of the time-frequency plane the DWT must be generalized.

Coefficient tree and subspace analysis tree for the **two-scale** (three levels) FWT filter bank. The subspace analysis tree provides a more compact representation of the decomposition in terms of subspaces.

\[
V_j = V_{j-1} \oplus W_{j-1}
\]
Wavelet Packets (cont…)

Subspace analysis tree and spectrum splitting for the \textbf{three-scale} (four levels) FWT filter bank. Here, we have three options for the expansion:

\begin{align*}
V_J &= V_{J-1} \oplus W_{J-1} \\
V_J &= V_{J-2} \oplus W_{J-2} \oplus W_{J-1} \\
V_J &= V_{J-3} \oplus W_{J-3} \oplus W_{J-2} \oplus W_{J-1}
\end{align*}

Wavelet Packets (cont…)

Wavelet packets are conventional wavelet transforms in which the details are filtered iteratively. For example, the three scale wavelet packet tree is

Indices \(A\) (approximation) and \(D\) (detail) denote the path from the parent to the node.
Wavelet Packets (cont…)

The cost of this generalization increases the computational complexity of the transform from $O(M)$ to $O(M \log M)$.

Also, the three scale packet almost triples the number of decompositions available from the three low pass bands.

C. Nikou – Digital Image Processing (E12)
Wavelet Packets (cont…)

The subspace of the signal may be expanded as

\[ V_J = V_{J-3} \oplus W_{-1} \oplus W_{-2,A} \oplus W_{-2,D} \oplus W_{-1,A} \oplus W_{-1,AD} \oplus W_{-1,DA} \oplus W_{-1,DD} \]

with spectrum:

It can also be expanded as

\[ V_J = V_{J-1} \oplus W_{J-1,A} \oplus W_{J-1,DA} \oplus W_{J-1,DD} \]

with spectrum:
Wavelet Packets (cont…)

In general, a $P$-scale 1-D wavelet packet (associated with $P+1$ level analysis trees) supports

$$D(P+1) = [D(P)]^2 + 1$$

unique decompositions, where $D(1)=1$ is the initial signal.

For instance, $D(4) = 26$ and $D(5) = 677$.

The problem increases dramatically in 2-D. A $P$-scale 2-D wavelet packet (associated with $P+1$ level analysis trees) supports

$$D(P+1) = [D(P)]^4 + 1$$

unique decompositions, where $D(1)=1$ is the initial signal.

For instance, $D(4) = 83.522$ possible decompositions. How do we select among them? Impractical to examine each one of them for a given application.
Wavelet Packets (cont…)

A common application of wavelets is image compression. Here we will examine a simplified example. We seek to compress the fingerprint image by selecting the “best” three-scale wavelet packet decomposition.

A criterion for the decomposition is the image energy

\[ E(f) = \sum_m \sum_n |f(m,n)|^2 \]

Algorithm
For each node of the analysis tree, from the root to the leaves:

**Step 1:** Compute the energy of the node \((E_p)\) and the energy of its offsprings \((E_A, E_H, E_V, E_D)\).

**Step 2:** If \(E_A + E_H + E_V + E_D < E_p\), include the offsprings in the analysis tree as they reduce the initial energy. Otherwise, keep only the parent which is a leaf of the tree.
Wavelet Packets (cont…)

Many of the 64 initial subbands are eliminated.

C. Nikou – Digital Image Processing (E12)