An Introductory Tutorial on Wavelets: Notes

ISPRS International Summer School, Advances in Medical Imaging

24–29 April, 2006, Aghios Nikolaos, Crete, Greece

Evangelos Roussos
Program in Applied and Computational Mathematics
Princeton University
eroussos@math.princeton.edu

1 Introduction

Wavelets are mathematical magnifying lenses: wavelet analysis is a tool for the *hierarchical decomposition* of functions. They integrate concepts from various scientific fields, like functional analysis, signal processing, statistics, etc.; see Fig. 1

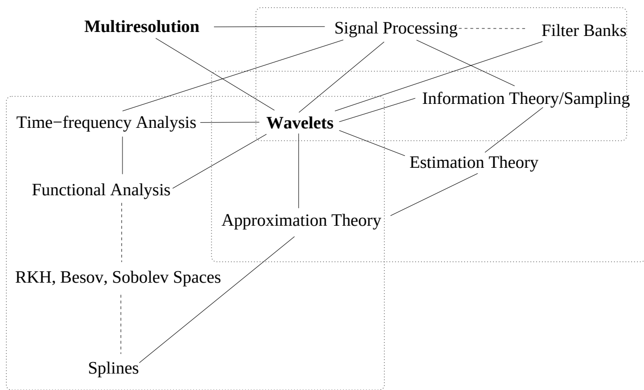

Figure 1: Wavelets analysis and allied fields.

2 Signals and Represention

2.1 Signals as functions and vectors

- Signals can be modelled as functions, $f : \Omega \to \mathcal{X}$, from the time– or space–domain to the space of amplitudes of the signal. Digital signals are from $\otimes = \mathbb{I} = \{(i_1, \ldots, i_d)\} \subseteq \mathbb{N}^d$, the space of d–tuples of integer indices (for example, $d = 2$ for images), to $\mathcal{X} \subseteq \mathbb{R}$.
- Functions can be thought of as vectors in a very high–dimensional space. Intuitively, we can understand this by discretising a function f with sampling interval Δt (Fig. 2.1), and letting $\Delta t \to 0$.

2.2 Domains, representations, and transforms

Physical, frequency, and wavelet domains.

- The physical (time– or space–) domain representation expresses a function f on $\Omega \subseteq \mathbb{R}^d$ as a combination of an impulse-train of Dirac δ–functions[1] at

[1] Strictly speaking, it should be called a 'distribution'.

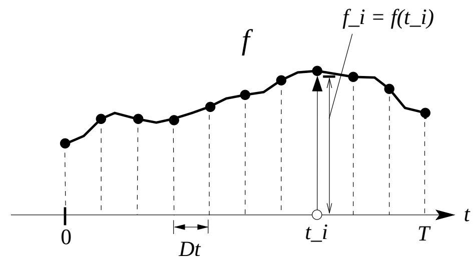

Figure 2: Sampling a continuous function f with sampling interval Δt.

time-points $t_i = i\Delta t$ for timeseries, spatial positions $(x_i, y_j) = (i\Delta x, j\Delta y)$ for images, etc; using $\mathbf{r}_i$ as a generic index,

$$f(\mathbf{r}) = \sum_i f_i \delta(\mathbf{r} - \mathbf{r}_i) = \sum_i f(\mathbf{r}_i)\delta(\mathbf{r} - \mathbf{r}_i), \quad \mathbf{r} \in \Omega \subseteq \mathbb{R}^d, \quad \mathbf{r}_i = i\Delta\mathbf{r}.$$

We can think of δ–function as "picking" the value of a function at each $\mathbf{r}_i$. The discrete equivalent is using the canonical basis, $\{\mathbf{e}_i\}_i$, $\mathbf{e}_i = (\ldots, 0, \ldots, 0, 1, 0, \ldots, 0, \ldots)$, with a 1 at position i.

- More generally, we can represent a signal as a sum (or integral) of bases, $\{\mathbf{b}_k\}$:

$$f(\mathbf{r}) = \sum_k c_k b_k(\mathbf{r}), \quad \mathbf{r} \in \Omega \subseteq \mathbb{R}^d, \tag{1}$$

  where d is the dimension of our space, for example $d = 2$ for images. *'Choosing a representation' means expressing our signal in a certain basis.* The bases are "prototypical signals", in a sense, and their amplitude is "modulated" by their corresponding coefficient.

- In a frequency domain (Fourier) representation, the basis functions are sinusoids, or complex exponentials, $\{e^{\iota\omega t}\}_\omega$, $\iota \stackrel{\text{def}}{=} \sqrt{-1}$:

$$f(t) = \frac{1}{2\pi}\int \mathrm{d}\omega \hat{f}(\omega) e^{\iota\omega t},$$

  where ω is the frequency; the representation is $f(t) \longmapsto \hat{f}(\omega)$, from t–space to ω–space. The Fourier domain is therefore useful for representing the *frequency content* of a signal.

- Fourier bases are perfectly localised w.r.t. frequency, ω, but their support[2] is the whole real axis, $(-\infty, \infty)$: *they are not localised in physical space.* This means that we cannot tell *when* a particular 'frequency event' happened.

[2] *'Support'* is the part of the domain of a function in which it is non–zero.

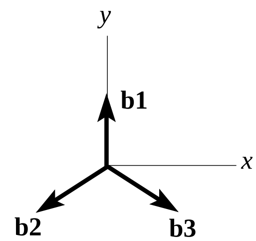

Figure 3: A 3-frame $\{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3\}$ in 2-D space, (x, y).

- Wavelets are basis functions that span the space of signals with finite energy[3], therefore they can represent any function in this space.

2.2.1 From bases to frames

- We can represent a vector with more "bases" than the dimensionality of the vector space itself (under some conditions): see Fig. 2.2.1.
- This is called an *overcomplete* representation, and it is very useful in many contexts. The elements of a frame are generically called *'atoms'*.

2.2.2 Localised Bases

Properties of localised bases bases: "chop up" signals into small pieces. Heisenberg boxes.

- If we combine a Fourier basis $e^{\iota\omega t}$ with a *window* $g(t)$ that has finite support, we cut off the part of the signal outside the window.
- By shifting the window, in physical space, by u, we get a family of atoms $\{g_{u,\omega}\}$:

$$g_{u,\omega}(t) \stackrel{\text{def}}{=} e^{\iota\omega t} g(t-u).$$

- This leads to the *windowed–* or *short–time* Fourier Transform (STFT), $f(t) \longmapsto \tilde{f}(u, \omega)$:

$$\tilde{f}(u,\omega) = \langle f, g_{u,\omega} \rangle = \int \mathrm{d}t f(t) g(t-u) e^{-\iota\omega t},$$

  where $\langle \cdot \rangle$ is the inner product.

[3]We can define the energy of a function by $\frac{1}{2}\int_0^{2\pi} \mathrm{d}x\, |f(x)|^2$.

2.2.3 Time–frequency Tiling

- We say that the atoms *'tile'* the time–frequency space.
  - We can visualise this by ploting the result of the transform in (u, ω)–space. The support of an atom localised at (u, ω) is

    $$\sigma_t \times \sigma_\omega = \left[u - \frac{\sigma_t}{2}, u + \frac{\sigma_t}{2}\right] \times \left[\omega - \frac{\sigma_\omega}{2}, \omega + \frac{\sigma_\omega}{2}\right];$$

    these are called 'Heisenberg boxes': *they represent the uncertainty, or trade–off, w.r.t. precise localisation in space versus frequency content.*
  - The Heisenberg boxes of δ functions are stripes with perfect localisation on the time axis, t, but infinite support on the frequancy axis.
  - Fourier bases have the exactly oposite representation: perfect localisation on the frequency axis, ω, but zero "resolution" on the time axis.
  - STFT–tiles are *identical* parallelograms, $\sigma_t \times \sigma_\omega$, shifted in time and space in order to cover the time–frequency plane.
- Wavelets are another kind of localised bases with very interesting properties, which attempt to balance time– and frequency–localisation; see next.

3 Wavelets

3.1 General properties

- Wavelets are functions that satisfy certain requirements:
  - They integrate to zero: this property makes them 'wave–s',
  - They are *well localised* in space, i.e. they have 'compact support' ('–lets').
- Wavelets also form families of "self-similar" atoms, in the sense that new atoms can be formed by the *dilation* and *translation* of a basic "template", so–called "mother", wavelet.

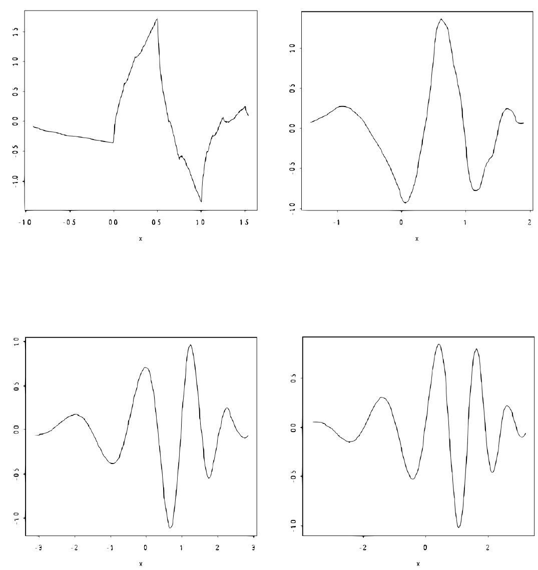

Figure 4: Members of the Daubechies class of wavelets [?]: from left–to–right, D2, D4, D8, D10. Each one can form a whole family of bases.

- The concept of *scale*: since wavelet bases are localised functions, and they can be formed by the dilation of a mother–wavelet, they naturally incorporate a notion of characteristic scale. Unlike Fourier bases, where the natural property is frequency, i.e. 'number of oscillations per unit length', wavelets can be (and usually are) constructed such that they are identically zero outside a certain range.
- Smoothness property: wavelets trade-off locality of support and smoothness: the less localised they are, the smoother they become.

  This is necessary, in order to have a mathematically stable construction.

These are extremely useful properties, since they allow us to "zoom" on a certain level of detail in the signal, and add more detail as needed.

There are many different types, depending on their other properties, like orthogonality, smoothness, locality of their support, their relation to equivalent digital filters, etc.; see Fig. 3.1.

3.2 Wavelet transform: Continuous

- As mentioned above, we can construct wavelets by translating and dilating a mother wavelet:

$$\psi_{s,\tau}(t) = \frac{1}{\sqrt{s}}\,\psi\left(\frac{t-\tau}{s}\right), \tag{2}$$

  where the factor $1/\sqrt{s}$ is for normalising the energy across scales.

- The generic wavelet transform of a function, $f(t) \longmapsto w(s,\tau)$ can then be written as

$$w(s,\tau) = \langle \psi^*_{s,\tau}, f \rangle = \int \mathrm{d}t\, \psi^*_{s,\tau}(t) f(t), \tag{3}$$

  where the star denotes complex conjugation. This is an inner product, or correlation, of our signal with the wavelet function. *It is a measure of how much the details of our signal at that particular scale and position "look like" our basis.*

- By taking all translations τ and dilations s of the mother wavelet, ψ, we get a very detailed picture of the information content of our signal, with respect to scale and spatial position. This is the *continuous wavelet transform*, CWT.

- A visual representation of the transform in (s,τ)–space is called a *scalogram.*

- We can also perform the inverse operation, $w(s,\tau) \longmapsto f(t)$, from wavelet space to physical space, by

$$f(t) = \int \mathrm{d}s\mathrm{d}\tau\, w(s,\tau)\psi_{s,\tau}(t). \tag{4}$$

3.3 Wavelet transform: Discrete

The Cohen–Daubechies–Feauveau (CDF) machinery (Fourier-based) for the construction of wavelets.

- The CWT is a redundant transform: in order to reconstruct the original signal from $\{w_{s,\tau}\}$ we do not need *all* dilations, s, and translations, τ.

  In fact, in many cases, for example in signal compression, we do *not* want to have all wavelet coefficients[4].

[4] Of course, in many others we *do* want to keep all the information.

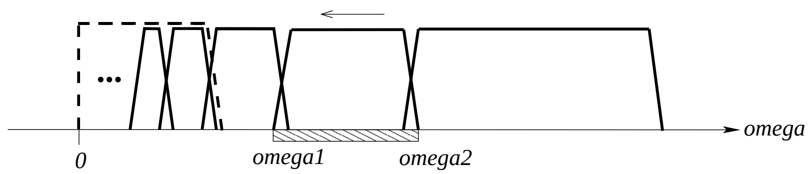

Figure 5: Wavelet spectra: a particular wavelet basis ψ_{jk} spans a region $[\omega_1, \omega_2]$ in the frequency domain. The scaling function, ϕ, covers the lower end of the frequency spectrum.

- It turns out that we can "sparsify" our set of "samples" $\{(s, \tau)\}$ in Eq. 3 to only a critical number of them $\{(s_j, \tau_k)\}$, and still be able to reconstruct the original signal[5]. We only need to translate and dilate at discrete steps:

$$\psi_{j,k}(t) = \frac{1}{\sqrt{{s_0}^j}}\, \psi\left(\frac{t - k\tau_0 {s_0}^j}{{s_0}^j}\right) \tag{5}$$

  We can choose $s_0 = 2$, which leads to *dyadic sampling*, i.e. sampling on a *logarithmic grid* of points $\{(s_j, \tau_k)\}$ in position–scale space.

- The discrete wavelet transform can therefore be written as

$$f(t) = \sum_{j,k} w_{j,k}\psi_{j,k}(t), \quad \forall j, k \in \mathbb{Z}, \quad t \in \mathbb{R}^d. \tag{6}$$

3.3.1 The View from the Fourier Domain

- Note that wavelets are signals too; therefore they have a Fourier representation.

  Now, recalling the fact from Fourier analysis that contraction (resp. dilation) by a in the physical domain causes a dilation (resp. contraction) by the same amount in the frequency domain, $\mathcal{F}[f(at)] = 1/|a|\,\mathcal{F}[\omega/a]$, $a \in \mathbb{R}$, and the fact that daughter–wavelets can be generated by scaling a mother–wavelet, we see that Eq. 6 amounts to adding a set of wavelet spectra in the Fourier domain, in order to capture the frequency content of f, as shown shematically in Fig. 3.3.1.

- Returning to the tiling idea: wavelets tile the time–frequency space by adaptively adjusting their width and height according to Eq. 5.

3.3.2 Multiresolution Analysis

Multiresolution Analysis (MRA) and Mallat's Pyramid Algorithm

[5]Technically: this reduced set still spans the space of signals with finite power.

- From Fig. 3.3.1 we see that a wavelet function has a band-like spectrum, spanning a region $[\omega_1, \omega_2]$ in the frequency domain. Therefore, reconstructing a signal f using Eq. 6 can be interpreted as filtering with a set of *band–pass* filters.

- Notice, though, that using this construction requires an infinite countable, in general, number of wavelet bases, since $|\Delta\omega| \to 0$, $\Delta\omega = \omega_2 - \omega_1$, as we keep dillating the wavelets[6].

  Stéphane Mallat, [?], made the ingenious observation that one could use another function instead, the *scaling function*, ϕ, with just the right frequency band. This then, combined with our wavelets, $\{\psi\}$ covers the whole spectrum of f. The action of the scaling function corresponds to *low–pass* filtering.

- Again, the scaling function can be represented in the Fourier domain by a linear combination of wavelet bases.

3.3.3 The Fast Wavelet Transform

Signal Processing: Filters for the implementation of the WT. Two–scale relations

- The remarkable thing is that the above analysis can be implemented extremely efficiently, via an algorithm with computational cost $\mathcal{O}(n)$.

- A particular wavelet family corresponds to, and is implemented with, a filter–bank. In particular, the original length–N signal is recursively passed through a pair of high–pass (**G**) and low–pass (**H**) filters. We store the result of **G**, which is a sequence of $N/2$ wavelet coefficients, and keep splitting and filtering the low–passed portion each time, until we reach sequences of unit length; one stop the algorithm at a higher level, of course.

- The above 'iterated filter bank' is the pyramid algorithm. Mathematically, it corresponds to set of relations between wavelets and scaling coefficients at two different scales, j and $j-1$.

- *The above construction implements an analysis of a signal into several levels of detail.*

[6] Intutively, an "infinitely small wavelet" would be "almost like" a Dirac δ and its spectrum would cover a very large region of the frequency domain. On the contrary an "infinitely large wavelet" would cover a very small region $\Delta\omega$ around a point ω_0.

4 Wavelets for Images

The above discussion, although quite generic, did not specifically focus on the issue of wavelets for two– or higher–dimensional signals.

A digital image can be represented as a matrix $\mathbf{A} = [A_{ij}] \in \mathbb{R}^{N \times N}$ with 'grey level' A_{ij} at pixel (i, j).

1. The application of the pyramid algorithm on an image can be done in 'sub–bands': first the rows of $\mathbf{A}$, $\{\mathbf{a}_i\}_i$, are filtered using $\mathbf{H}$ and $\mathbf{G}$, giving two new matrices $\mathbf{H}_r\mathbf{A}$ and $\mathbf{G}_r\mathbf{A}$ of dimensions $N \times N/2$.

2. The same operation is performed on the columns of each of these two matrices giving four matrices, $\mathbf{H}_c\mathbf{H}_r\mathbf{A}$, $\mathbf{G}_c\mathbf{H}_r\mathbf{A}$, $\mathbf{H}_c\mathbf{G}_r\mathbf{A}$, and $\mathbf{G}_c\mathbf{G}_r\mathbf{A}$, of dimension $N/2 \times N/2$.

   The matrix $\mathbf{H}_c\mathbf{H}_r\mathbf{A}$ contains the low–pass filtered image and the rest the high–pass filtered ones.

3. Store $\{\mathbf{G}_c\mathbf{H}_r\mathbf{A}, \mathbf{H}_c\mathbf{G}_r\mathbf{A}, \mathbf{G}_c\mathbf{G}_r\mathbf{A}\}$ and continue filtering the low–pass image until we get a matrix of dimensions 1×1, or stop at a higher node of the pyramid.

The above algorithmic construction corresponds to the application of the wavelet transform with two-dimensional wavelets that are *tensor products* of one–dimensional ones.

4.1 Applications:

I. Compression/Sparse Representation,

II. Denoising.