Wavelet analysis : from NMR spectroscopy to motion estimation

Jean-Pierre Antoine

Institut de Recherche en Mathématique et Physique, Université catholique de Louvain
B-1348 Louvain-la-Neuve, Belgium
E-mail: jean-pierre.antoine@uclouvain.be

Abstract

We review the general properties of the wavelet transform, both in its continuous and its discrete versions, in one or two dimensions, and we describe some of its applications in signal processing. We also consider its extension to two dimensions, that is, for image analysis, to more general, curved, manifolds and to the space-time context, for the analysis of moving objects.

1 What is wavelet analysis?

Wavelet analysis is a particular time- or space-scale representation of signals which has found a wide range of applications in physics, signal processing and applied mathematics in the last few years. In order to get a feeling for it and to understand its success, let us consider first the case of one-dimensional signals.

The traditional tool in signal processing is Fourier transform

$$s(x) \quad \leftrightarrow \quad \widehat{s}(\xi) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-i\xi x} s(x)\, dx$$

Strictly speaking, it applies to *stationary* signals only. However, it is a fact that most real life signals are nonstationary. They often contain transient components, sometimes very significant physically, and mostly cover a wide range of frequencies. In addition, there is frequently a direct correlation between the characteristic frequency of a given segment of the signal and the time duration of that segment. Low frequency pieces tend to last a long interval, whereas high frequencies occur in general for a short moment only. Human speech signals are typical in this respect. Vowels have a relatively low mean frequency and last quite long, whereas consonants

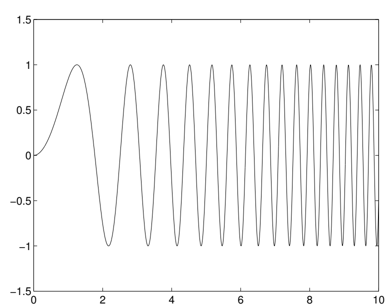


Figure 1: A nonstationary signal (chirp)

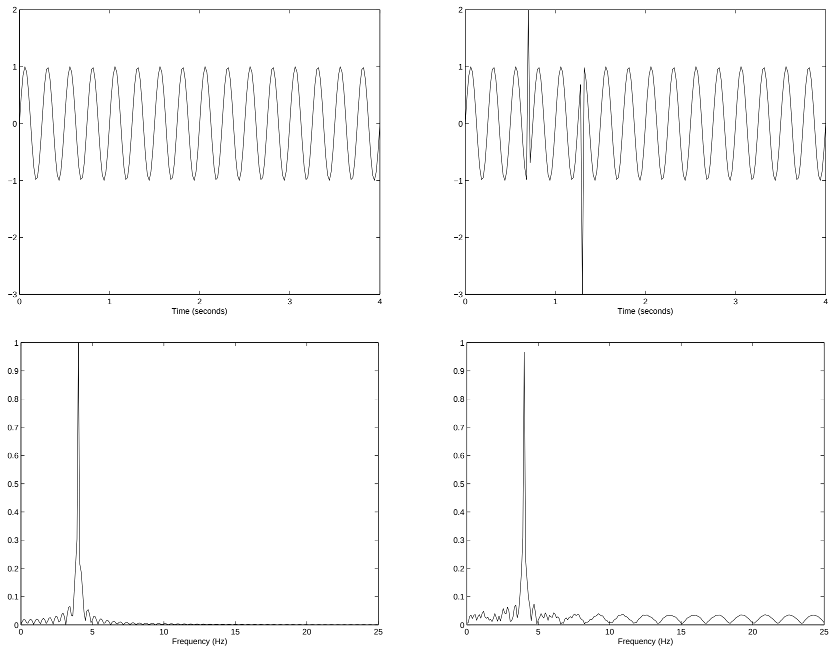


Figure 2: (Top) A pure sine wave; the same with two delta perturbations added; (Bottom) The respective Fourier transforms

contain a wide spectrum, up to very high frequencies, especially in the attack, but they are very short.

Clearly standard Fourier analysis is inadequate for treating such signals, since it loses all information about the time localization of a given frequency component. In addition, it is very uneconomical. If a segment of the signal is almost flat, thus carries little information, one still has to sum an (alternating) infinite series for reproducing it. Worse yet, Fourier analysis is highly unstable with respect to perturbation, because of its *global* character. For instance, if one adds an extra term, with a very small amplitude, to a linear superposition of sine waves, the signal will barely be modified, but the Fourier spectrum will be completely perturbed (Fig.2). Such pathologies do not occur if the signal is represented in terms of *localized* components.

Therefore, signal analysts turn to *time-frequency* (TF) representations. The idea is that one needs *two* parameters. One, called a, characterizes the frequency, the other one, b, indicates the position in the signal. This concept of a TF representation is in fact quite old and familiar. The most obvious example is simply a musical score (Figure 3)!

If one requires, in addition, the transform to be *linear*, a general TF transform will take the

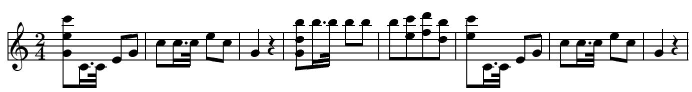


Figure 3: A traditional time-frequency representation of a signal (from Mozart's Don Giovanni, Act 1).

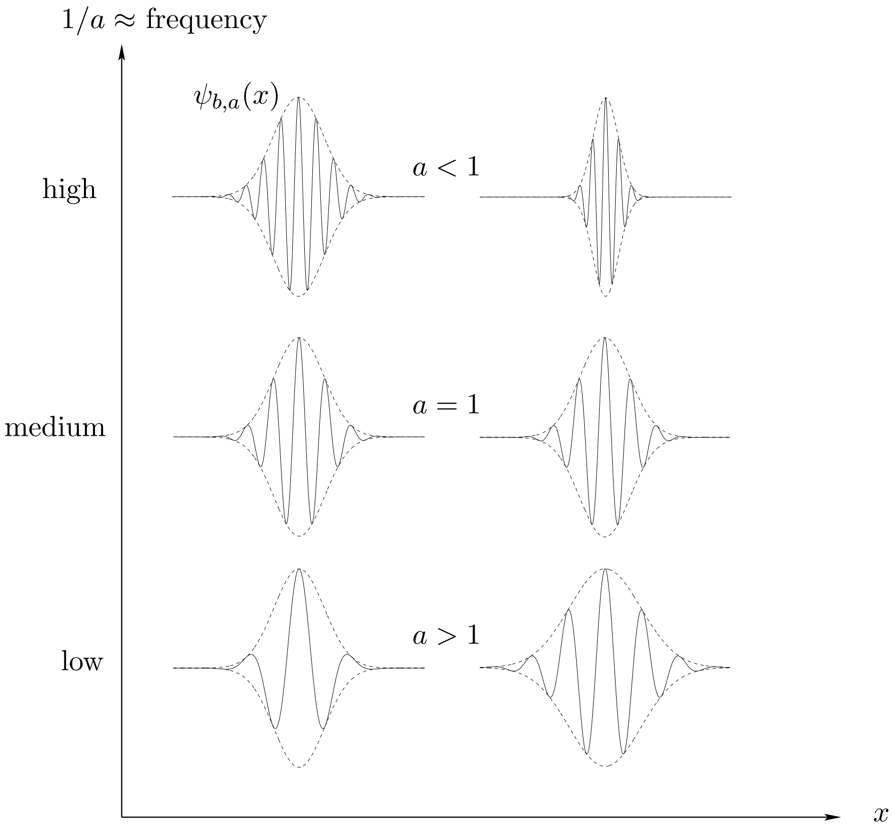


Figure 4: The function $\psi_{b,a}(x)$ for different values of the scale parameter a : in the case of the Windowed Fourier Transform (left); in the case of the wavelet transform (right).

form:

$$s(x) \mapsto S(b,a) = \int_{-\infty}^{\infty} \overline{\psi_{b,a}(x)}\, s(x)\, dx, \tag{1.1}$$

where s is the signal and $\psi_{b,a}$ the analyzing function (we denote the time variable by x, in view of the extension to higher dimensions). The assumption of linearity is actually nontrivial, for there exists a whole class of quadratic or, more properly, sesquilinear time-frequency representations. The prototype is the so-called Wigner-Ville transform, introduced originally by E.P. Wigner [27] in quantum mechanics (in 1932!) and extended by J. Ville [26] to signal analysis:

$$W_s(b,\xi) = \int_{-\infty}^{+\infty} e^{-i\xi x}\, \overline{s\Big(b - \frac{x}{2}\Big)}\, s\Big(b + \frac{x}{2}\Big)\, dx\,, \quad \xi = 1/a. \tag{1.2}$$

Notice that, contrary to the linear version (1.1), this transform is intrinsic, it contains no external probe ψ, which unavoidably influences the result.

Within the class of linear TF transforms, two stand out as particularly simple and efficient, the Windowed or Short-Time Fourier Transform (STFT)[1] and the wavelet transform (WT). For both of them, the analyzing function $\psi_{b,a}$ is obtained by a group action on a basic (or mother) function ψ, only the group differs. The essential difference between the two is in the way the

[1]Sometimes called the *Gabor transform*, although Gabor considered only a discretized version of it [14].

frequency parameter a is introduced. In both cases, b is simply a time translation. The kernels of the two transforms can be written as follows.

1. For the *Windowed Fourier Transform*, one uses
$$\psi_{b,a}(t) = e^{it/a}\,\psi(t-b). \tag{1.3}$$
Here ψ is a window function and the a-dependence is a modulation ($1/a \sim$ frequency); the window has constant width, but the lower a, the larger the number of oscillations in the window.

2. For the *wavelet transform*, instead, one takes
$$\psi_{b,a}(t) = \frac{1}{\sqrt{a}}\,\psi\Big(\frac{t-b}{a}\Big). \tag{1.4}$$
The action of a on the function ψ is a dilation ($a > 1$) or a contraction ($a < 1$). The shape of the function is unchanged, it is simply spread out or squeezed. In particular, the effective support of ψ_{ba} varies as a function of a.

The difference between the two transforms is illustrated on Figure 4. In the sequel we will discuss only the WT.

Finally a word about references. As a first contact, introductory articles such as [21] or [15] may be a good suggestion, followed by the popular book of B. Burke Hubbard [16] or the elementary book of Y. Meyer [20] As for textbooks, we note the volumes of B. Torrésani [25] (in French!), I. Daubechies [13] or S. Mallat [18]. Our main source of information, however, especially concerning the applications of the CWT in physics, will be van den Berg's survey [10] and our own monographs [1, 2]. In addition, we ought to mention that most of the figures have been obtained with our wavelet toolbox [29].

Before proceeding, it is worth emphasizing the fact that there are actually *three* different stages of the wavelet transform.

1. *The Continuous WT (CWT)*
$$S(b,a) = |a|^{-1/2} \int_{-\infty}^{\infty} \overline{\psi\left(\frac{x-b}{a}\right)}\, s(x)\, dx, \quad a \neq 0,\, b \in \mathbb{R}.$$
Here one considers *all* values of a and b (often restricted to $a > 0$); this is useful for feature detection.

2. *The discretized CWT:* discretization is needed for numerical implementation, but this requires a suitable choice of the sampling grid. In this case, one cannot get orthonormal bases, only *frames*, thus a redundant representation.

3. *The discrete WT (DWT):* here one chooses a preselected grid (mostly dyadic) and one may obtain (bi)orthonormal bases from multiresolution analysis. This approach is good for *data compression* in the presence of large amounts of data.

We will discuss all three in the sequel. We emphasize the fact that the (discretized) CWT is incompatible with the DWT, they correspond to totally different philosophies. We may also note the following analogy:

CWT	$\Longleftrightarrow$	Fourier integral
discretized CWT	$\Longleftrightarrow$	Fourier series
DWT	$\Longleftrightarrow$	discrete FT.

2 The one-dimensional CWT

2.1 Basic formulas

As announced above, the basic formulas read, in time domain and in frequency domain, respectively,

$$
\begin{aligned}
S(b,a) &= \langle \psi_{b,a} | s \rangle \\
&= a^{-1/2} \int_{-\infty}^{\infty} \overline{\psi\left(a^{-1}(x-b)\right)}\, s(x)\, dx && (2.1) \\
&= a^{1/2} \int_{-\infty}^{\infty} \overline{\widehat{\psi}(a\xi)}\, \widehat{s}(\xi)\, e^{i\xi b}\, d\xi, && (2.2)
\end{aligned}
$$

where $a \neq 0$ is a scale parameter, $b \in \mathbb{R}$ is a translation parameter and the hat denotes a Fourier transform. Thus the pair (b, a) runs over the time-scale half-plane $\mathbb{R}^2_+$. In practice one uses also the restriction $a \neq 0$, which is physically more natural.

In these relations, s is a square integrable function and the analyzing wavelet ψ assumed to be well localized *both* in the space (or time) domain and in the frequency domain. Here the minimal requirement is that $\psi \in L^1 \cap L^2$, but in practice stronger conditions are usually imposed (like $\psi \in \mathcal{S}$, Schwartz's space of fast decreasing functions). In consequence, the CWT provides good bandpass filtering both in x *and* in ξ.

Moreover, ψ must satisfy the following admissibility condition, which guarantees the invertibility of the WT (see (2.12) below):

$$c_\psi := 2\pi \int_{-\infty}^{\infty} d\xi\, \frac{|\widehat{\psi}(\xi)|^2}{|\xi|} < \infty. \tag{2.3}$$

In most cases, this condition may be reduced to the requirement that ψ has zero mean (the condition is necessary, and become sufficient upon a slight restriction on ψ):

$$\widehat{\psi}(0) = 0 \iff \int_{-\infty}^{\infty} \psi(x)\, dx = 0. \tag{2.4}$$

Note : if one takes $a > 0$, one needs a slightly different admissibility condition :

$$c_\psi := 2\pi \int_0^{\infty} \frac{|\widehat{\psi}(\xi)|^2}{|\xi|}\, d\xi = 2\pi \int_{-\infty}^{0} \frac{|\widehat{\psi}(\xi)|^2}{|\xi|}\, d\xi < \infty$$

(the equality is automatic if ψ is real).

In addition, ψ is often required to have a certain number of *vanishing moments:*

$$\int_{-\infty}^{\infty} x^n\, \psi(x)\, dx = 0,\ n = 0, 1, \ldots N. \tag{2.5}$$

This property improves the efficiency of ψ at detecting singularities in the signal, since it is blind to polynomials up to order N.

Finally, it is often useful, but not essential, that ψ be *progressive*, which means that $\widehat{\psi}$ is real and $\widehat{\psi}(\xi) = 0$ for $\xi < 0$ (in the language of signal processing, ψ is an analytic signal).

In order to fix ideas, we exhibit here two commonly used wavelets (see Figure 5).

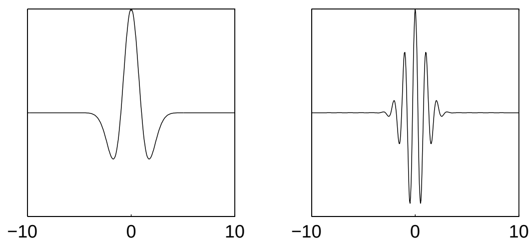


Figure 5: (left) The Mexican hat or Marr wavelet; (right) The real part of the Morlet wavelet, for $\xi_o = 5.6$.

(1) The Mexican hat or Marr wavelet

This wavelet is simply the second derivative of a Gaussian:

$$\psi_H(x) = (1 - x^2)\, e^{-x^2/2}, \quad \widehat{\psi}_H(\xi) = \xi^2\, e^{-\xi^2/2}. \tag{2.6}$$

It is real and admissible, it has 2 vanishing moments $n = 0, 1$, but it is not progressive.

(2) The Morlet wavelet

This wavelet is essentially a plane wave within a Gaussian window:

$$\psi_M(x) = e^{ik_o x}\, e^{-x^2/2\sigma_o^2} + h(x), \quad \widehat{\psi}_M(\xi) = \sigma_o\, e^{-[(\xi-\xi_o)\sigma_o]^2/2} + \widehat{h}(\xi). \tag{2.7}$$

It is complex, but not progressive, strictly speaking (numerically it is). Here the correction term h must be added in order to satisfy the admissibility condition (2.4), but in practice one will arrange that this term be numerically negligible ($\leqslant 10^{-4}$) and thus can be omitted (it suffices to choose the basic frequency $|\xi_o|$ large enough, namely, one has to take $|\sigma_o \xi_o| > 5.5$).

These two wavelets have very different properties and, naturally, they will be used in quite different situations. Typically, the Mexican hat is sensitive to singularities in the signal, and it yields a genuine time-scale analysis. On the other hand, since the Morlet wavelet is complex, it will catch the phase of the signal, hence will be sensitive to frequencies, and will lead to a time-frequency analysis, somewhat closer to a Gabor analysis. In both cases, additional flexibility is obtained by adding a width parameter to the Gaussian.

2.2 Localization properties and interpretation

We must now make more precise the localization conditions on the wavelet ψ. Assume that the numerical support of $\psi(x)$ is an interval of length L around 0, and that of $\widehat{\psi}(\xi)$ is an interval of length Ξ, centered around the mean frequency ξ_o (by numerical support, we mean the smallest set outside of which the function is numerically negligible). Then the numerical support of $\psi_{b,a}(x)$ is an interval of length aL around b, while that of $\widehat{\psi}(\xi)$ is an interval of length Ξ/a, centered around ξ_o/a.

Therefore (see Figure 6):

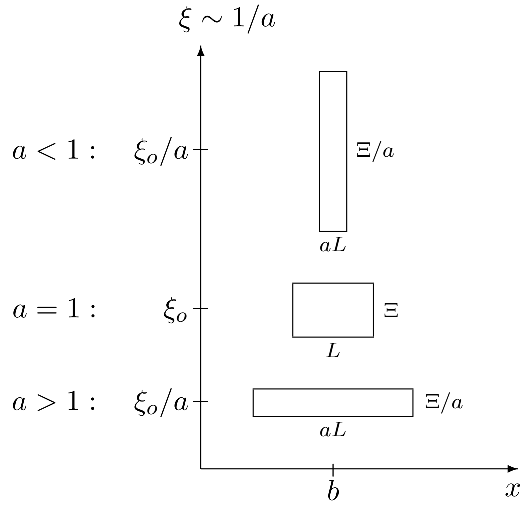


Figure 6: Support properties of $\psi_{b,a}$ and $\widehat{\psi_{b,a}}$.

- if $a \gg 1$, $\psi_{b,a}$ is a wide window (long duration) and $\widehat{\psi_{b,a}}$ is peaked around the small frequency ξ_o/a: for large scales, the CWT is sensitive to *low frequencies*, and thus yields a rough analysis.
- if $a \ll 1$, $\psi_{b,a}$ is a narrow window (short duration), and $\widehat{\psi_{b,a}}$ is wide and centered around the high frequency ξ_o/a: for very small scales, the CWT is sensitive to *high frequencies* (small details).

Combining now these localization properties with the fact that the CWT is a convolution with a zero mean filter, we conclude that:

- the CWT provides a local filtering in time (b) and scale (a):

$$S(b,a) \not\approx 0 \quad \Longleftrightarrow \quad \psi_{b,a}(x) \approx s(x).$$

- the CWT may be interpreted as a mathematical microscope, with optics ψ, position b, and magnification $1/a$.
- the CWT works at constant relative bandwidth : $\Delta\xi/\xi = \text{const.}$

As a consequence, one may consider the CWT as a *singularity detector and analyzer*.

2.3 Mathematical properties

Given an admissible wavelet ψ, the CWT $W_\psi : s(x) \mapsto S(b,a)$ is a linear map, with the following properties:

(1) *Covariance* under translation and dilation:

$$W_\psi : s(x - x_o) \mapsto S(b - x_o, a) \tag{2.8}$$

$$W_\psi : \frac{1}{\sqrt{a_o}}\, s\big(\frac{x}{a_o}\big) \mapsto S\big(\frac{b}{a_o}, \frac{a}{a_o}\big). \tag{2.9}$$

(2) *Energy conservation :*

$$\int_{-\infty}^{\infty} |s(x)|^2 dx \, = c_\psi^{-1} \, \iint_{\mathbb{R}^2_+} \, |S(b,a)|^2 \frac{da\, db}{a^2}. \tag{2.10}$$

Thus, $|S(b,a)|^2$ may be interpreted as the energy density in the time-scale half-plane.

The relation (2.10) means that W_ψ is an *isometry* from the space of signals $L^2(\mathbb{R})$ onto a *closed* subspace $\mathcal{H}_\psi$ of $L^2(\mathbb{R}^2_+, da\, db/a^2)$, namely, the space of wavelet transforms. An equivalent statement is that the wavelet ψ generates a *resolution of the identity:*

$$c_\psi^{-1} \, \iint_{\mathbb{R}^2_+} \, |\psi_{b,a}\rangle \, \langle \psi_{b,a}| \, \frac{da\, db}{a^2} = I. \tag{2.11}$$

(3) *Reconstruction formula :* As a consequence, W_ψ is invertible on its range $\mathcal{H}_\psi$ by the adjoint map, i.e., we obtain an exact reconstruction formula:

$$s(x) = c_\psi^{-1} \, \iint_{\mathbb{R}^2_+} \, \psi_{b,a}(x) \, S(b,a) \, \frac{da\, db}{a^2}. \tag{2.12}$$

In other words, we have a representation of the signal $s(x)$ as a linear superposition of wavelets $\psi_{b,a}$ with coefficients $S(b,a)$.

(4) The projection $P_\psi : L^2(\mathbb{R}^2_+, da\, db/a^2) \to \mathcal{H}_\psi$ is an integral operator, with kernel

$$K(b', a'; b, a) = c_\psi^{-1} \, \langle \psi_{b',a'} | \psi_{b,a} \rangle. \tag{2.13}$$

The kernel K is the autocorrelation function of ψ and it is a *reproducing kernel*. Indeed, the function $f \in L^2(\mathbb{R}^2_+, da\, db/a^2)$ is the WT of a certain signal iff it satisfies the *reproduction property*

$$f(b', a) = c_\psi^{-1} \, \iint_{\mathbb{R}^2_+} \, \langle \psi_{b',a'} | \psi_{b,a} \rangle \, f(b,a) \, \frac{da\, db}{a^2} \, . \tag{2.14}$$

This means that the CWT is a highly redundant representation!

The last statement implies that the full information must be contained is a small subset of half-plane. This property may be exploited in two ways: either one takes a discrete subset, which leads to the theory of *frames*, or one considers only the lines of local maxima, called *ridges*.

2.4 Applications of the 1-D CWT

The main usage of the 1-D CWT consists in analyzing transient phenomena, detecting abrupt changes in a signal or comparing it with a given pattern. here is a list of significant applications. References to the original papers may be found in our textbook [1].

- *Sound and acoustics*
  musical synthesis, speech analysis, disentangling of an underwater refracted wave.

- *Geophysics, meteorology*
  analysis of various long time series of geophysical origin: gravimetry, geomagnetism, astronomy (Solar physics), analysis of meteorological radar data ...

- *Fractals*
  artificial (diffusion limited aggregates) or natural (arborescent growth phenomena) fractals.

- *Turbulence, fluid mechanics*
  identification of coherent structures and hierarchical structure, detection of intermittency.

- *Atomic physics*
  analysis of high order harmonic generation, controlled emission of ultrashort light pulses, in the 10^{-18}s (attosecond) range.

- *Spectroscopy*
  subtraction of unwanted spectral lines or filtering out of background noise, e.g. in NMR spectroscopy (see Section 2.5 below).

- *Analysis of local singularities*
  detection of singularities in a signal, fine characterization of their strengths.

- *Quantum cosmology*
  wavelet quantization regularizes the gravitational singularity in a Friedmann-Lemaître universe.

- *Edge detection and shape characterization*
  take contour of an object as a complex curve in the plane, then apply CWT.

- *Medical and biological applications*
  analyzing or monitoring various electrical or mechanical phenomena in the brain (EEG, VEP), the heart (ECG) or the visual system, statistical analysis of correlations in DNA sequences.

- *Engineering and applied science*
  monitoring of nuclear, electrical or mechanical installations (detection of anomalies).

2.5 Application to NMR spectroscopy

The physical phenomenon underlying NMR spectroscopy may be described as follows [1, Sec. 12.7.1]. When a sample is placed in a static magnetic field, nuclei with a magnetic moment align along this applied field, resulting in a net magnetization. In its equilibrium state, the magnetization is static and does not induce a signal in the receiver antenna. In order to obtain information, one must first excite the nuclei with a radio frequency pulse. After such a pulse, the magnetizations precess around the static field at angular frequencies characteristic of their chemical environment and relax to their equilibrium state. This precession induces a signal in the receiver antenna. The signal to be analyzed is the Fourier transform (spectrum) of the damped response curve of the protons (this is called an FID signal, for Free Induction Decay). It contains a large number of narrow peaks, the spectral lines, corresponding to different metabolites, but many among them are useless, coming for instance from the protons of the solvent. These peaks, which may be quite big, must be subtracted, and the position and amplitude of the relevant ones measured with precision. In addition, the spectra are often quite noisy, and must be 'cleaned' before any useful measurement can be performed.

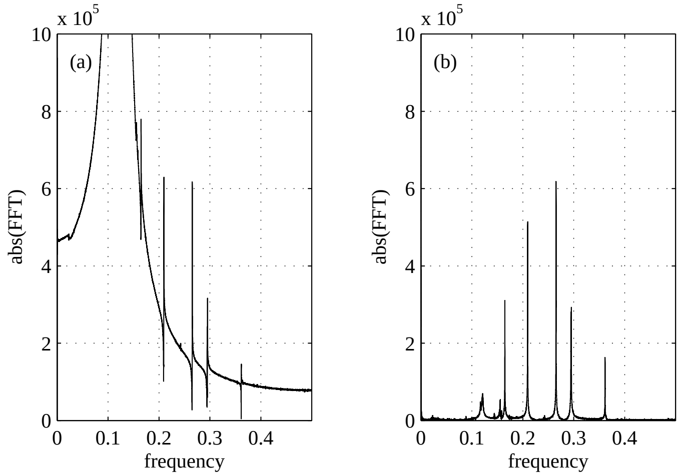


Figure 7: Application of the CWT in NMR spectroscopy: Subtraction of an unwanted peak. (a) The original spectrum; (b) The spectrum reconstructed after subtraction of the water peak.

We present here three typical applications on the CWT in this context. The first one is an example of peak subtraction, Figure 7. The original spectrum (left) exhibits a huge parasite peak, due to the protons of the solvent (water), that masks to a large extent the interesting structures. The analysis consists in isolating this peak, on the CWT, subtracting it from the spectrum and reconstructing the remaining part. The result (right) is a spectrum where all the fine details are now clearly visible, and have not been perturbed by the removal of the large peak.

Figure 8 is an example of noise filtering. The original signal (top) consists of a number of damped sinusoids embedded in noise. The dominant peaks are localized with help of the CWT ridge algorithm, the remnant of the spectrum is subtracted and the filtered spectrum is reconstructed, using one of the reconstruction formulas.

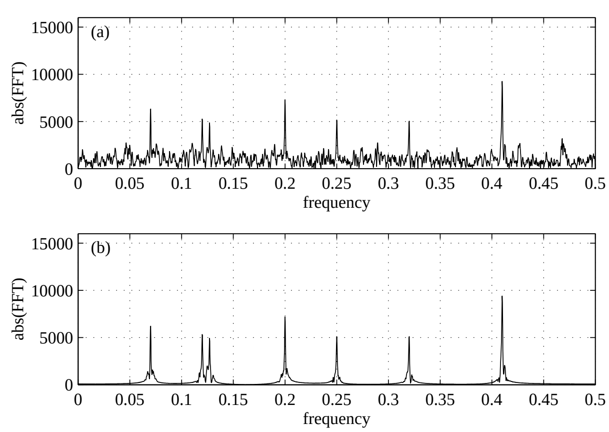


Figure 8: Noise filtering in NMR spectroscopy: (a) The original spectrum; (b) The spectrum reconstructed after noise removal.

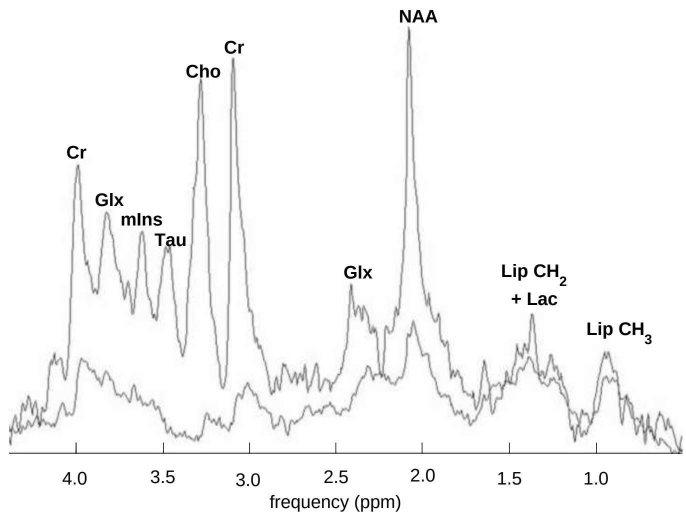


Figure 9: *In vivo* short-echo time MRS spectrum acquired at 4.7T in a mouse brain, superimposed on a measured *in vivo* macromolecular spectrum.

However, in an *in vivo* spectrum, the analysis and quantification of the metabolite contributions is difficult, because of the low signal to noise ratio, the number of overlapping frequencies and the contamination of the signal of interest with water and a baseline originating from macromolecules and lipids. A typical example is shown in Figure 9. A CWT with the Morlet wavelet can be used for quantifying signals acquired *in vivo*, that is, recovering parameters such as metabolite amplitudes, frequencies and damping factors [23]. Although the Morlet wavelet is quite efficient for obtaining such results, it is not optimal for each kind of spectrum. The problem is that the Morlet wavelet is standard, but it is not directly related to the analyzed signal. An alternate method, presented in [17] is to *define* wavelet functions directly from the MRS data themselves, by the autocorrelation function of the metabolite profiles, In this way, one obtains wavelets, called *metabolite-based wavelets*. More precisely, given a metabolite X, one acquires its profile $s(X)$. Then the autocorrelation function of $s(X)$, properly averaged to zero, may be taken as the wavelet ψ_X adapted to the metabolite X.

A spectacular application of this concept is a technique for determining the presence of a given metabolite in a composite signal, which contains an unspecified mixture. This proceeds as follows. Given a composed signal, one analyzes it with the wavelet ψ_X. Then, if the CWT of $s(X)$ has a strong horizontal ridge at $a = 1$, this means that X is present in the mixture. Otherwise X is *not* present in the mixture. Figure 10 shows a concrete example. The distinguished metabolite is called NAA (see Figure 9). We show in Figure 10 the analysis of two composite signals. The left panels present, on top, the CWT of the composite signal using the NAA-wavelet, and, below, the corresponding ridges (lines of maximum modulus). In the same way, the right panels present the CWT and its ridges for the composite signal, but with the NAA contribution removed. Indeed, we see a strong ridge at scale $a = 1$ when NAA is present (left panels), but that ridge is no longer present if the NAA contribution is removed from the signal.

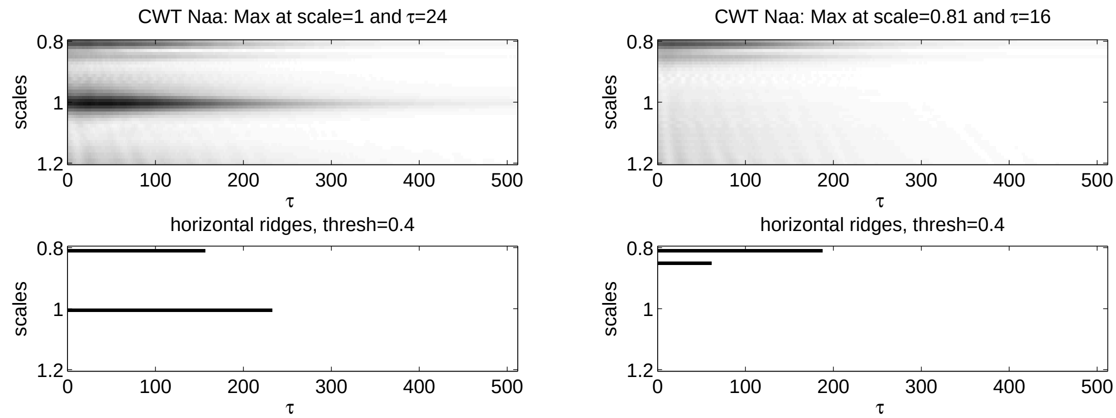


Figure 10: CWT using the *in vitro* NAA-wavelet on (a) the composite signal and (b) the composite signal lacking the NAA contribution.

This technique exploits optimally the properties of wavelet analysis, whose philosophy is always to exploit a maximum of *a priori* information. It could prove very useful in practical implementations of MRs spectroscopy, in particular for clinical applications. For a fully detailed exposition, we refer to the report [24]

2.6 Discretization of the CWT

Clearly, the CWT must be discretized for numerical implementation, but then the question arises of choosing an adequate sampling grid. We will say that a discrete lattice $\Gamma = \{a_j, b_{j,k}, j, k \in \mathbb{Z}\}$ yields good discretization if the following *exact* relation holds:

$$s = \sum_{j,k\in\mathbb{Z}} \langle \psi_{jk}, s \rangle \widetilde{\psi}_{jk}\,, \tag{2.15}$$

with $\psi_{jk} \equiv \psi_{b_{j,k},a_j}$ and where $\widetilde{\psi}_{jk}$ is explicitly constructible from ψ_{jk}. A common choice is the so-called *dyadic* grid $a_j = 2^{-j}$, $b_{j,k} = k \cdot 2^{-j}$, which yields

$$\psi_{jk}(x) = 2^{j/2}\psi(2^j\, x - k), \quad j, k \in \mathbb{Z}\,. \tag{2.16}$$

The problem is that approach usually leads to frames, not bases. It is worth recalling here the definition. A *frame* in the Hilbert space $\mathcal{H}$ is a family $\{\psi_{jk}\}$ for which there exists two constants $\mathsf{m} > 0, \mathsf{M} < \infty$ such that

$$\mathsf{m}\,\|s\|^2 \;\leqslant\; \sum_{j,k\,\in\mathbb{Z}} |\langle\psi_{jk}|s\rangle|^2 \;\leqslant\; \mathsf{M}\,\|s\|^2,\ \forall\, s \in L^2(\mathcal{R}). \tag{2.17}$$

The upper bound means that the map $s \mapsto \{\langle\psi_{jk}|s\rangle\}$ is continuous from L^2 to l^2, whereas the lower bound guarantees the numerical stability of the inverse map [13]. The constants m,M are called *frame bounds*. If $\mathsf{m} = \mathsf{M} \neq 1$, the frame is said to be *tight*. Of course, if $\mathsf{m} = \mathsf{M} = 1$ and $\|\psi_{jk}\| = 1,\ \forall\, j, k$, we simply get an orthonormal basis. A detailed analysis then shows that the expansion (2.15) converges essentially as a power series in the quantity $|\mathsf{M}/\mathsf{m} - 1|$. We note that both the Mexican hat and the Morlet wavelet yield good, but non tight, frames.

3 The discrete WT (DWT)

One of the successes of the WT was the discovery that it is possible to construct functions ψ for which $\{\psi_{jk}, j, k \in \mathbb{Z}\}$ is indeed an orthonormal basis of $L^2(\mathbb{R})$, while keeping the good properties of wavelets, including space *and* frequency localization. In addition, this approach yields fast algorithms, and this is the key to the usefulness of wavelets in many applications.

The construction is based on two facts. First, almost all examples of orthonormal bases of wavelets can be derived from a multiresolution analysis, and then the whole construction may be transcripted into the language of Quadrature Mirror Filters (QMF), familiar in the signal processing literature.

A *multiresolution analysis* of $L^2(\mathbb{R})$ is an increasing sequence of closed subspaces

$$\ldots \subset V_{-2} \subset V_{-1} \subset V_0 \subset V_1 \subset V_2 \subset \ldots$$

with $\bigcap_{j\in\mathbb{Z}} V_j = \{0\}$ and $\bigcup_{j\in\mathbb{Z}} V_j$ dense in $L^2(\mathbb{R})$, and such that

(1) $f(x) \in V_j \Leftrightarrow f(2x) \in V_{j+1}$;

(2) There exists a function $\phi \in V_0$, called a *scaling function*, such that the family $\{\phi(x-k), k \in \mathbb{Z}\}$ is an orthonormal basis of V_0.

Combining conditions (1) and (2), one sees that $\{\phi_{jk}(x) := 2^{j/2}\phi(2^j x - k),\ k \in \mathbb{Z}\}$ is an orthonormal basis of V_j. The space V_j can be interpreted as an *approximation* space at resolution 2^j. Defining now

$$V_j \oplus W_j = V_{j+1}, \tag{3.1}$$

we see that W_j contains the additional *details* needed to improve the resolution from 2^j to 2^{j+1}. Thus one gets the decomposition

$$L^2(\mathbb{R}) = \bigoplus_{j\in\mathbb{Z}} W_j. \tag{3.2}$$

Equivalently, one may choose a lowest resolution level j_o and obtain

$$L^2(\mathbb{R}) = V_{j_o} \oplus \left(\bigoplus_{j=j_o}^{\infty} W_j \right), \tag{3.3}$$

The crucial theorem then asserts the existence of a function ψ, sometimes called the *mother wavelet*, explicitly computable from ϕ, such that $\{\psi_{jk}(x) := 2^{j/2}\psi(2^j x - k), j \in \mathbb{Z}\}$ is an orthonormal basis of W_j. Then $\{\psi_{jk}(x) := 2^{j/2}\psi(2^j x - k),\ j, k \in \mathbb{Z}\}$ constitutes an orthonormal basis of $L^2(\mathbb{R})$: these are the *orthonormal wavelets*. Well-known examples include the Haar wavelets, the B-splines, and the various Daubechies wavelets.

In practice, one starts from a sampled signal, taken in some V_J, and then the decomposition (3.3)) is replaced by the finite representation

$$V_J = V_{j_o} \oplus \left(\bigoplus_{j=j_o}^{J-1} W_j \right). \tag{3.4}$$

Now a natural question is to decide which version should be used in a concrete case, the (discretized) CWT or the DWT? As usual in a wavelet context, the answer depends on the application at hand:

- for feature detection, the CWT is preferable, since no *a priori* choice for is made for a, b ; the CWT is more flexible and also more robust to noise, but only frames will be available in general.
- for large amount of data or data compression, one should use the DWT; it yields bases, it is faster, but it is also more rigid.

This explains why a number of generalizations have been introduced in order to alleviate these drawbacks of the DWT, such as biorthogonal wavelets, wavelet packets, continuous wavelet packets (integrated wavelets), redundant WT (on a rectangular lattice) or "Second generation" wavelets (the so-called lifting scheme). For all these, we refer to the textbooks quoted in the bibliography.

4 The two-dimensional CWT

4.1 Basic formulas

Now we turn to the two-dimensional CWT, which has become a major tool in image processing. In this context, an image is a two-dimensional signal of finite energy, represented by a complex-valued function $s \in L^2(\mathbb{R}^2, d^2\vec{x})$:

$$\|s\|^2 = \int_{\mathbb{R}^2} d^2\vec{x}\, |s(\vec{x})|^2 < \infty. \tag{4.1}$$

Given an image s, all the geometric operations we want to apply to it are obtained by combining three elementary transformations of the plane, namely, rigid translations in the plane of the image, dilations or scaling (global zooming in and out) and rotations. Explicitly, the transformations act on $\vec{x} \in \mathbb{R}^2$ in the familiar way:

$$\begin{array}{ll} \text{(i)} & \text{translation by } \vec{b} \in \mathbb{R}^2 : \vec{x} \mapsto \vec{x}\,' = \vec{x} + \vec{b}, \\ \text{(ii)} & \text{dilation by a factor } a > 0 : \vec{x} \mapsto \vec{x}\,' = a\vec{x}, \\ \text{(iii)} & \text{rotation by an angle } \theta : \vec{x} \mapsto \vec{x}\,' = r_\theta(\vec{x}), \end{array}$$

where

$$r_\theta \equiv \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix},\ 0 \leqslant \theta < 2\pi,$$

is the familiar 2×2 rotation matrix.

The combined action of these three types of transformations is realized by the following unitary map in the space $L^2(\mathbb{R}^2, d^2\vec{x})$ of finite energy signals:

$$\left[U(\vec{b}, a, \theta)s\right](\vec{x}) \equiv s_{\vec{b},a,\theta}(\vec{x}) = a^{-1} s(a^{-1}\, r_{-\theta}(\vec{x} - \vec{b})). \tag{4.2}$$

In terms of this action, the basic formulas for the 2-D CWT read

$$S(\vec{b}, a, \theta) = \langle \psi_{\vec{b},a,\theta} | s \rangle \tag{4.3}$$

$$= a^{-1} \int_{\mathbb{R}^2} \overline{\psi(a^{-1}\, r_{-\theta}(\vec{x} - \vec{b}))}\, s(\vec{x})\ d^2\vec{x} \tag{4.4}$$

$$= a \int_{\mathbb{R}^2} e^{i\vec{b}\cdot\vec{k}}\, \overline{\widehat{\psi}(ar_{-\theta}(\vec{k}))}\, \widehat{s}(\vec{k})\ d^2\vec{k}. \tag{4.5}$$

As in 1-D, we have to impose an *admissibility* condition on the wavelet ψ, namely,

$$c_\psi := (2\pi)^2 \int_{\mathbb{R}^2} \frac{|\widehat{\psi}(\vec{k})|^2}{|\vec{k}|^2}\, d^2\vec{k} < \infty, \tag{4.6}$$

which again may be replaced in practice by the following necessary condition:

$$\widehat{\psi}(\vec{0}) = 0 \iff \int_{\mathbb{R}^2} d^2\vec{x}\ \psi(\vec{x}) = 0. \tag{4.7}$$

The important observation to make here is that all the formulas are almost identical in 1-D and in 2-D ! As a consequence, the interpretation of the CWT as a singularity analyzer still holds, and the mathematical properties of the 2-D transform strictly parallel those of its 1-D counterpart.

4.2 Group-theoretical justification

However, before proceeding, we should pause and ask the reason of this situation. As can be expected, the answer lies in group theory.

(1) In one dimension

Observe that dilations and translations are *affine* transformations of the line:

$$x = (b, a)y = ay + b, \quad a \neq 0,\ b \in \mathbb{R},$$

and they obey the following composition rule: $(b, a)(b', a') = (b + ab', aa')$. Thus, the set of all these affine transformations constitutes a group, the affine group of the line, $\{(b, a)\} := G_{\rm aff} \simeq \mathbb{R}^2_*$.

The natural action of (b, a) on a signal, $\psi \mapsto U(b, a)\psi$, reads

$$[U(b, a)\psi](x) = a^{-1/2}\psi\left(\frac{x - b}{a}\right) \tag{4.8}$$

and U is a unitary irreducible representation of $G_{\rm aff}$ in $L^2(\mathbb{R})$. In addition, U is *square integrable*, namely,

$$\psi \text{ admissible } \iff \iint_{G^+_{\rm aff}} |\langle U(b, a)\psi|\psi\rangle|^2\, \frac{db\, da}{a^2} < \infty. \tag{4.9}$$

This explains the mathematical meaning of the admissibility condition of wavelets.

Restricting to $a > 0$, one gets the *connected* affine group $G^+_{\rm aff}$ (or $ax + b$ group) and U is a UIR of it in $L^2(\mathbb{R}^+)$.

(2) In two dimensions

In two dimensions, dilations, translations, and rotations together constitute the *similitude group* of the plane: $\mathrm{SIM}(2) = \mathbb{R}^2 \rtimes \mathbb{R}^+_* \times \mathrm{SO}(2))$, with action

$$\vec{x} = (\vec{b}, a, \theta)\vec{y} = ar_\theta\vec{y} + \vec{b}.$$

This transformation is realized by the following unitary map on finite energy signals:

$$\big[U(\vec{b}, a, \theta)s\big](\vec{x}) = a^{-1}s(a^{-1}\, r_{-\theta}(\vec{x} - \vec{b})) \tag{4.10}$$

and U is a unitary irreducible representation of SIM(2) in $L^2(\mathbb{R}^2)$. In addition, U is *square integrable*:

$$\psi \text{ admissible } \Longleftrightarrow \iiint_{\mathrm{SIM}(2)} \left|\langle U(\vec{b}, a, \theta)\psi|\psi\rangle\right|^2 \, d^2\vec{b} \, \frac{da}{a^3} \, d\theta \; < \infty \tag{4.11}$$

4.3 Mathematical properties of the 2-D CWT

For the same reason, the mathematical properties of the 2-D CWT are essentially the same as in the 1-D case, so we list them without further comment:

(1) *Energy conservation*

$$c_\psi^{-1} \iiint_{\mathrm{SIM}(2)} |S(\vec{b}, a, \theta)|^2 \, d^2\vec{b} \, \frac{da}{a^3} \, d\theta = \int_{\mathcal{R}^2} |s(\vec{x})|^2 \, d^2\vec{x}, \tag{4.12}$$

i.e., the 2-D CWT is an isometry from the space of signals $L^2(\mathbb{R}^2)$ onto a closed subspace of $L^2(\mathrm{SIM}(2))$, the space of wavelet transforms.

(2) *Reconstruction formula*
Inverting the CWT by the adjoint map, we get

$$s(\vec{x}) = c_\psi^{-1} \iiint_{\mathrm{SIM}(2)} \psi_{\vec{b},a,\theta}(\vec{x}) \, S(\vec{b}, a, \theta) \, d^2\vec{b} \, \frac{da}{a^3} \, d\theta, \tag{4.13}$$

i.e., we have a decomposition of the signal in terms of the analyzing wavelets $\psi_{\vec{b},a,\theta}$, with coefficients $S(\vec{b}, a, \theta)$.

(3) *Reproduction property (reproducing kernel)*

$$S(\vec{b}', a', \theta') = c_\psi^{-1} \iiint d^2\vec{b} \, \frac{da}{a^3} \, d\theta \, \langle \psi_{\vec{b}',a',\theta'} | \psi_{\vec{b},a,\theta} \rangle \, S(\vec{b}, a, \theta). \tag{4.14}$$

(4) the CWT is *covariant* under translations, dilations and rotations, that is, under SIM(2): the correspondence $W_\psi : s(\vec{x}) \mapsto S(\vec{b}, a, \theta)$ implies the following ones

$$\begin{aligned} s(\vec{x} - \vec{b}_o) &\;\mapsto\; S(\vec{b} - \vec{b}_o, a, \theta), \\ a_o^{-1} s(a_o^{-1}\vec{x}) &\;\mapsto\; S(a_o^{-1}\vec{b}, a_o^{-1} a\theta), \\ s(r_{\theta_o}(\vec{x})) &\;\mapsto\; S(r_{-\theta_o}(\vec{b}), a, \theta - \theta_o). \end{aligned}$$

Note that translation covariance (“shift invariance”) is lost in the standard formulation of the discrete WT, based on multiresolution, which creates problems in pattern recognition, for instance.

4.4 Localization properties and interpretation

Assume

- ψ and $\widehat{\psi}$ as well localized as possible (compatible with Fourier uncertainty property)
- num supp $\psi(\vec{x})$ $\sim$ ‘disk’ of diameter T, centered around $\vec{0}$
  num supp $\widehat{\psi}(\vec{k})$ $\sim$ ‘disk’ of diameter Ω, centered around $\vec{k}_o$

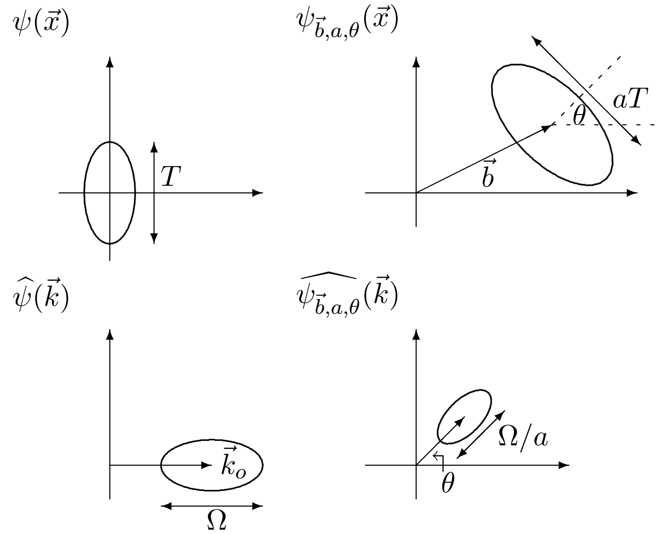


Figure 11: Support properties under the basic operations ($a = 1.5, \theta = 45°$); (Top) In the time domain; (Bottom) In the frequency domain.

Then

$$\begin{aligned} \text{num supp } \psi_{\vec{b},a,\theta} &\sim \text{ 'disk' of diameter} \simeq aT \text{ centered around } \vec{b} \text{ and rotated by an angle } \theta \\ \text{num supp } \widehat{\psi_{\vec{b},a,\theta}} &\sim \text{ 'disk' of diameter} \simeq \Omega/a, \text{ centered around } r_\theta(\vec{k}_o)/a \text{ and rotated by } \theta \end{aligned}$$

Note that the product of the two diameters is constant (it has to be bounded below by a fixed constant, by Fourier's theorem).

As a consequence,

- if $a \gg 1, \psi_{\vec{b},a,\theta}$ is a wide window and $\widehat{\psi_{\vec{b},a,\theta}}$ is peaked around a small spatial frequency $r_\theta(\vec{k}_o)/a$, thus it is sensitive to *low spatial frequencies* (rough analysis);

- if $a \ll 1, \psi_{\vec{b},a,\theta}$ is a narrow window and $\widehat{\psi_{\vec{b},a,\theta}}$ is wide and centered around a high spatial frequency $r_\theta(\vec{k}_o)/a$, hence it has a good spatial localization and it is sensitive to *high spatial frequencies* (small details).

The 2-D formulas being the exact analogues of the ones in 1-D, in particular the admissibility condition (4.6) or (4.7), clearly the interpretation of the CWT will be exactly the same as in 1-D. Thus, combining the localization properties of ψ with the fact that the CWT is a convolution with a zero mean function, we conclude again that the 2-D CWT

- yields a local filtering, this time in all three variables $\vec{b}, a, \theta$, in particular, a *directional* filtering;

- may be seen as a mathematical directional microscope with optics ψ, global magnification $1/a$, and orientation tuning parameter θ;

- works with constant *relative* bandwidth: $\Delta k/k = \text{const}, \quad k = |\vec{k}|$, thus it is particularly efficient for large spatial frequencies or small scales; thus it provides a good detection of *discontinuities* in images, such as point singularities (contours, corners) or directional features (edges, segments).

Therefore we may say that the 2-D CWT is a detector and analyzer of singularities (edges, contours, corners, ...).

Besides the previous requirements, one can improve the efficiency of the wavelets by imposing additional properties, such as restrictions on the support of ψ and of $\widehat{\psi}$, or *vanishing moments*, up to order $N \geqslant 1$ ($N = 0$ yields the admissibility condition) :

$$\int_{\mathbb{R}^2} x^\alpha \, y^\beta \, \psi(\vec{x}) \, d^2\vec{x} \; = 0, \quad \vec{x} = (x, y), \quad 0 \leqslant \alpha + \beta \leqslant N.$$

This will improve the efficiency at detecting singularities in the signal, since the transform is then blind to smoothest part of the signal, i.e., a polynomial of degree up to N, which contains little information, in general. For instance, if $N = 1$, the transform erases any linear *trend* in the signal, such as a linear gradient of luminosity.

4.5 Choice of the analyzing wavelet

Even more than in 1-D, it is important to choose a wavelet that is well adapted to the problem at hand. One can distinguish two classes, isotropic wavelets and direction sensitive wavelets.

(i) *Isotropic wavelets*

If one decides to perform a pointwise analysis, or if directions are irrelevant, it is sufficient to use a rotation invariant wavelet. Two standard examples are:

(1) *The 2-D Mexican hat wavelet*

This wavelet (originally introduced by Marr in his pioneering work on vision [19]) is simply the Laplacian of a Gaussian :

$$\psi_H(\vec{x}) = (2 - |\vec{x}|^2) \, \exp(-\tfrac{1}{2}|\vec{x}|^2), \quad \widehat{\psi_H}(\vec{k}) = |\vec{k}|^2 \, \exp(-\tfrac{1}{2}|\vec{k}|^2) \tag{4.15}$$

(2) *The Difference-of-Gaussians or DOG wavelet*

$$\psi_D(\vec{x}) = \tfrac{1}{2\alpha^2} \; \exp(-\tfrac{1}{2\alpha^2}|\vec{x}|^2) - \exp(-\tfrac{1}{2}|\vec{x}|^2) \qquad (0 < \alpha < 1) \tag{4.16}$$

This is a very good substitute to the Mexican hat; for $\alpha^{-1} = 1.6$, there are almost undistinguishable.

(ii) *Directional wavelets*

On the other hand, if the goal is to detect directional features in an image, or to perform directional filtering, clearly one should resort to a direction sensitive wavelet. The most efficient solution is a *directional* wavelet, that is, a wavelet $\psi(\vec{x})$ such that the numerical support of its Fourier transform $\widehat{\psi}(\vec{k})$ is contained in a convex cone with apex at the origin. Two useful examples are as follows.

(1) *The 2-D Morlet wavelet*

$$\psi_M(\vec{x}) = \exp(i\vec{k}_o \cdot \vec{x}) \exp(-\tfrac{1}{2}|\vec{x}|^2) + \text{corr.}$$
$$\widehat{\psi_M}(\vec{k}) = \exp(-\tfrac{1}{2}|\vec{k} - \vec{k}_o|^2) + \text{corr.}$$

As in 1-D, the correction term is required in order to satisfy the admissibility condition, and here too, it is negligible for $|\vec{k}_o| \geqslant 5.6$. The 2-D Morlet wavelet is shown in Figure 12 for $\epsilon = 2, \vec{k}_o = (0, 6)$.

(2) *Conical wavelets*, with support in the convex cone

$$\mathcal{C}(-\alpha, \alpha) \equiv \{\vec{k} \in \mathcal{R}^2 \,|\, -\alpha \leqslant \arg \vec{k} \leqslant \alpha,\ \alpha < \pi/2\}$$

$$\widehat{\psi}_m^{\mathrm{GCM}} = \begin{cases} (\vec{k} \cdot \vec{e}_{-\tilde{\alpha}})^m (\vec{k} \cdot \vec{e}_{\tilde{\alpha}})^m \ \exp(-\tfrac{1}{2}k_x^2),\ \vec{k} \in \mathcal{C}(-\alpha, \alpha) \\ 0,\ \text{otherwise} \end{cases} \tag{4.17}$$

where $\vec{e}_\phi$ is the unit vector in the direction ϕ and $\tilde{\alpha} = -\alpha + \pi/2$, so that $\vec{e}_{-\alpha} \cdot \vec{e}_{\tilde{\alpha}} = \vec{e}_\alpha \cdot \vec{e}_{-\tilde{\alpha}} = 0$. The frequency space version of this so-called Gaussian conical wavelet is shown in Figure 12, in the case $m = 4, \alpha = 10°$.

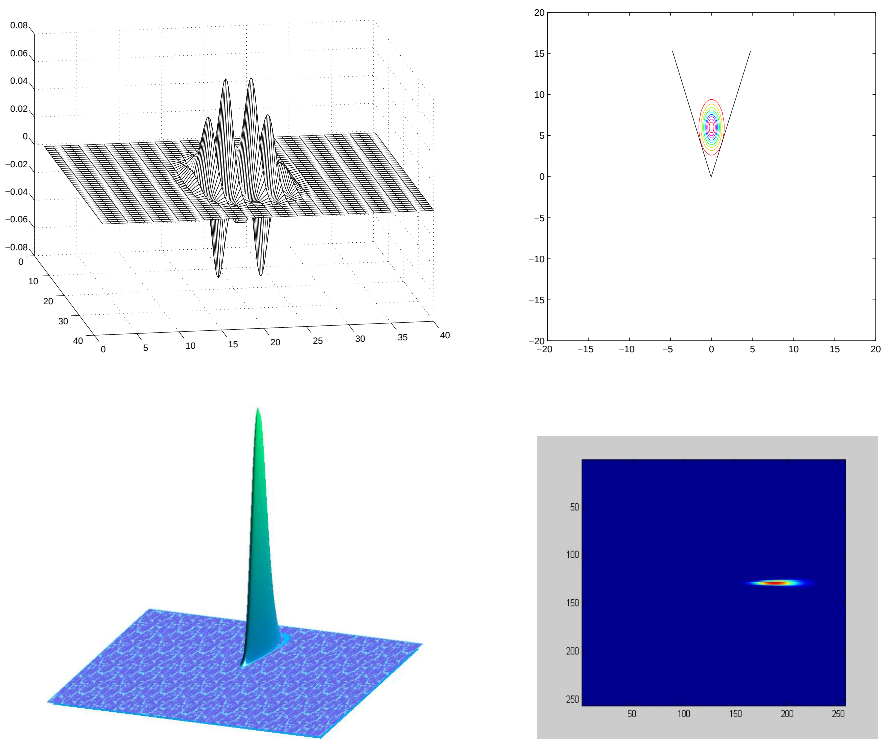


Figure 12: Two directional wavelets, in position space (Left) and in spatial frequency space (Right): (Top) The 2-D Morlet wavelet; (Bottom) The Gaussian conical wavelet

4.6 Applications of the 2-D CWT

As in 1-D, even more so, the 2-D CWT has been applied to a considerable number of problems. It is useful to distinguish between two types of applications, namely those pertaining to image processing and those belonging to genuine physical problems. Detailed information and references may be found in [2].

(a) Image processing

- *Image denoising*
  removal of noise in images using directional wavelets.

- *Contour detection, character recognition* detection of edges, contours, corners ...

- *Object detection and recognition in noisy images* automatic target recognition (ATR), application to infrared radar imagery, using both position and scale-angle features..

- *Image retrieval* recognition of a particular image in a large data bank, characterization of images by particular features.

- *Medical imaging* Magnetic resonance imaging (MRI), contrast enhancement, segmentation.

- *Detection of symmetries in 2-D patterns* detection of discrete inflation (rotation + dilation) symmetries, quasicrystals (mathematical and genuine), quasiperiodic point sets.

- *Watermarking of images:* adding a robust, but invisible, signature in images (e.g. with directional wavelets).

(b) Physical applications

- *Astronomy and astrophysics* structure of the Universe, cosmic microwave background (CMB) radiation, feature detection in images of the Sun, detection of gamma-ray sources in the Universe.

- *Geophysics* fault detection in geology, detection of arrival time of individual seismic waves in seismology, hurricane detection in climatology.

- *Fluid dynamics* detection of coherent structures in turbulent fluids, measurement of a velocity field, disentangling of an underwater acoustic wave train.

- *Fractals and the thermodynamical formalism* analysis of 2-D fractals by the WTMM method, determination of fractal dimension, unraveling of universal laws, shape recognition and classification of patterns.

- *Texture analysis* classification of textures, "Shape from texture" problem.

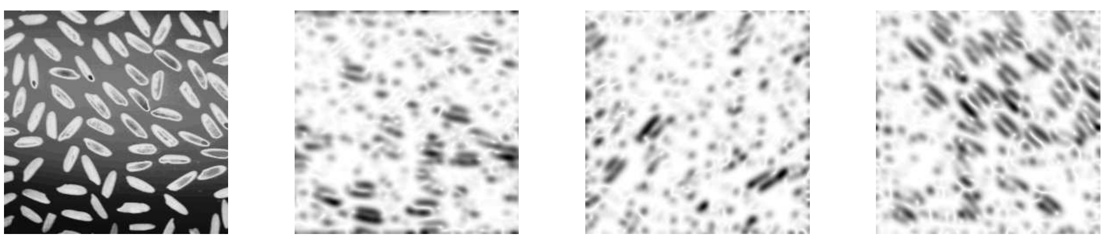
- *Fluid dynamics*
  detection of coherent structures in turbulent fluids, measurement of a velocity field, disentangling of an underwater acoustic wave train.

(a) (b) (c)

Figure 13: Directional filtering of the 'Mikado" pattern with a Conical wavelet oriented at $\theta = 0°$: (a) The original pattern; (b) Its CWT; (c) The CWT after thresholding t 25%

We give two examples of of physical applications. The first application is directional filtering, a nontrivial problem with the standard methods of image processing, in two different setups. The first one is a synthetic object, a bunch of segments (as in the "Mikado" game). Fig. 13 shows successively the signal, its CWT obtained with a (Cauchy) conical wavelet and the latter after thresholding.

The second example is a real image representing bacteria. Fig. 14 shows successively the signal, then the same filtered in there different directions, with a Gaussian conical wavelet. The panel (d) shows that there is a preferred direction at 135°. In addition, there is a density gradient from left to right.



(a) (b) (c) (d)

Figure 14: Directional filtering of the 'bacteria" pattern: (a) The original signal; (b) The same filteried at $-10°$; (c) The same at $45°$; (d) The same at $135°$

The second example id in the field of fluid dynamics. The aim is to measure the velocity field of a 2-D turbulent flow around an obstacle. Velocity vectors are materialized by small segments, by the technique of discontinuous tracers. Tiny plastic balls are seeded into the flow and illuminated by a 'plane of light', in order to get a 2-D image. Then two successive photos are taken with a fast CCD camera, with exposure times of 700 μs and 6000 μs, respectively. In this way one gets a 'dot-bar' signature for each tracer, which materializes the direction and the length of the local velocity (see an example in Figure 15, taken from [28]). In order to get sufficiently many data points, one superposes several such pictures, typically 16. First one computes the WT of the resulting image with a Morlet wavelet, which selects those vectors that are closely aligned with the wavelet. Then a second analysis is performed with a wavelet oriented in the orthogonal direction, thus completely misoriented with respect to the selected vectors. Now the WT sees only the tips of the vectors and their length may be easily measured.

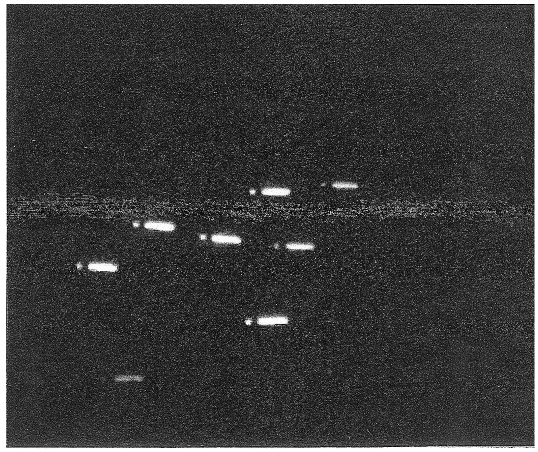
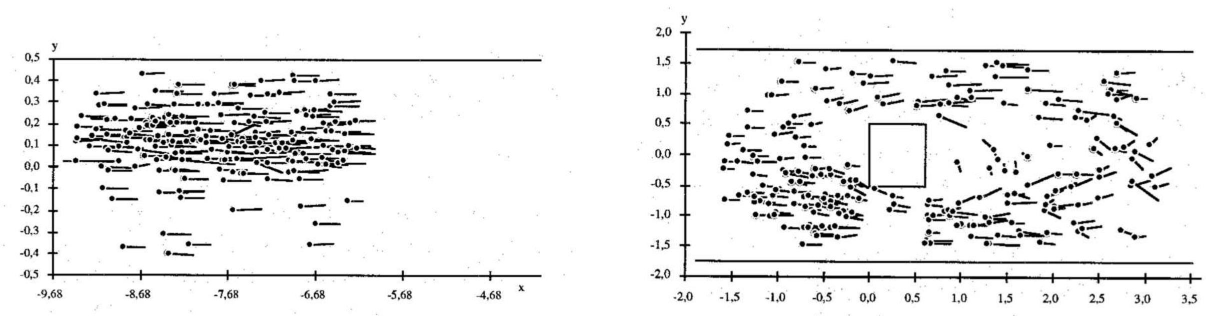


Figure 15: The dot-bar signature of tracers in the fluid flowing from left to right (from [28]).

The same two operations are then repeated with various successive orientations of the wavelet. Using appropriate thresholdings, the complete velocity field may thus be obtained, in a totally automated fashion, with an efficiency sensibly better than with more traditional methods. Two examples of reconstructed velocity fields from [28] are given in Figure 16, corresponding to a quasi-laminar flow and a turbulent flow around an obstacle (again the flow comes from the left; units are normalized to the size of the experimental area).



(a) (b)

Figure 16: Two examples of reconstructed velocity fields: (a) A quasi-laminar flow; (b) A turbulent flow around an obstacle (from [28]).

5 Extending the CWT to curved manifolds

5.1 Generalities

Many situations in physics yield data on *non-flat manifolds*:

- *sphere* : geophysics, cosmology (CMB), statistics, ...
- *two-sheeted hyperboloid* : cosmology (an open expanding model of the universe), optics (catadioptric image processing, where a sensor overlooks a hyperbolic mirror)

- *paraboloid* : optics (catadioptric image processing)
- *torus* : nuclear fusion (Tokamak), loop quantum gravity

How can one find suitable analysis tools?

A possible solution is to extend the *continuous* wavelet transform. Translation of the wavelet is easy, by an isometry of the manifold, i.e., an element of SO(3), SO(1,2)... the CWT is a local transform, with locality controlled by a dilation (to be defined!). However, in practice, the usual CWT works with discrete frames hence one needs also discrete wavelet frames on the manifold. An alternative solution is to extend the *discrete* wavelet transform.

Let us look first at the general problem. Given a manifold $\mathcal{M}$, how to derive a CWT on it? First, one should identify the operations to be performed on the finite energy signals living on $\mathcal{M}$, i.e., functions in $L^2(\mathcal{M}, d\mu)$, with μ a suitable measure on $\mathcal{M}$, namely,

(i) Motions are provided by the isometries of $\mathcal{M}$;

(ii) Dilations *on* $\mathcal{M}$ (zoom in/out) should form one-parameter abelian group $A \sim \mathcal{R}_+$.

If $\mathcal{M}$ admits a global isometry group $G_{\rm iso}$, merge it with the dilation group A into a global group G. Next, find a unitary irreducible, square integrable representation U of G in $L^2(\mathcal{M}, d\mu)$ and write (we assume that μ is G-invariant)

$$\psi_g(\zeta) := [U(g)\psi](\zeta) = \psi(g^{-1}\zeta),\ g \in G, \zeta \in \mathcal{M}$$

Then the CWT of $f \in L^2(\mathcal{M}, d\mu)$ with respect to the (admissible) wavelet ψ is defined as

$$W_\psi f(g) := \langle \psi_g | f \rangle = \int_{\mathcal{M}} \overline{\psi(g^{-1}\zeta)}\, f(\zeta)\, d\mu(\zeta),\ g \in G\,.$$

The corresponding *reconstruction formula* is then given by

$$f(\zeta) = c(\psi)^{-1/2} \int_G W_\psi f(g)\, \psi_g(\zeta)\, d\mu_\ell(g)\,,$$

where ψ is admissible iff $c(\psi) < \infty$ and μ_ℓ is the left Haar measure on G.

This is the scheme we have followed for the 1-D and the 2-D.

(1) *CWT in 1-D*

- $\mathcal{M} = \mathbb{R},\ G = G^+_{\rm aff} =$ "$ax + b$" group (connected affine group of the line)
- $\psi_{b,a}(x) = a^{-1/2}\psi\big(a^{-1}(x-b)\big),\ g = (b, a),\ b \in \mathbb{R}, a > 0$
- $\widehat{\psi_{b,a}}(k) = a^{1/2}\, \widehat{\psi}(ak)\, e^{-ibk}$

(2) *CWT in 2-D*

- $\mathcal{M} = \mathbb{R}^2,\ G = \mathsf{SIM}(2) = \mathbb{R}^2 \rtimes (\mathbb{R}^+_* \times \mathrm{SO}(2))$ (similitude group)
- $\psi_{\vec{b},a,\theta}(x) = a^{-1}\psi\big(a^{-1}r_{-\theta}(\vec{x} - \vec{b})\big),\ g = (\vec{b}, a, \theta),\ \vec{b} \in \mathbb{R}^2, a > 0, \theta \in [0, 2\pi),$
  $r_\theta = 2 \times 2$ rotation matrix
- $\widehat{\psi_{\vec{b},a,\theta}}(\vec{k}) = a\, \widehat{\psi}(ar^{-\theta}\vec{k})\ e^{-i\vec{b}.\vec{k}}$

The same approach extends to higher dimensions as well.

(3) *CWT in 3-D*

- $\mathcal{M} = \mathbb{R}^3,\ G = \mathrm{SIM}(3) = \mathbb{R}^3 \rtimes (\mathbb{R}^+_* \times \mathrm{SO}(3))$ (similitude group)
- $\psi_{\vec{b},a,\rho}(x) = a^{-3/2}\psi\big(a^{-1}\rho^{-1}(\vec{x}-\vec{b})\big),\ \vec{b}\in\mathbb{R}^3, a>0, \rho\in\mathrm{SO}(3)$
- $\widehat{\psi}_{\vec{b},a,\rho}(\vec{k}) = a^{3/2}\,\widehat{\psi}(a\rho^{-1}\vec{k})\ e^{-i\vec{b}.\vec{k}}$

  If ψ is axisymmetric, one may replace SO(3) by $\mathbb{S}^2 = \mathrm{SO}(3)/\mathrm{SO}(2)$, so that the wavelet is indexed by $(\vec{b}, a, \omega)$ instead of $(\vec{b}, a, \rho)$.

(3) *More complicated cases*

- Two-sphere $\mathbb{S}^2$, $G_{\rm iso} = \mathrm{SO}(3)$
- Two-sheeted hyperboloid $\mathbb{H}^2$, $G_{\rm iso} = \mathrm{SO}(2,1)$

These cases will be discussed in the next sections.

5.2 The WT on the two-sphere

(1) *The spherical CWT*

A CWT on the two-sphere $\mathbb{S}^2$ may be obtained by lifting the plane 2-D CWT from the tangent plane at North Pole onto $\mathbb{S}^2$ by inverse *stereographic projection* [4, 5, 8]. The construction relies on (nontrivial) group-theoretical considerations: embedding the isometry group SO(3) of $\mathbb{S}^2$ into its conformal group $\mathrm{SO}_0(3,1)$. It yields a full spherical CWT, with correct Euclidean limit, i.e., the spherical CWT on a sphere of radius R tends to the usual 2-D CWT in the tangent plane when $R \to \infty$ (here one uses tyne technique of group contraction). In addition, one may construct spherical frames, both semi-continuous (only the scale is discrete) and fully discrete, but one needs a generalized notion of frame (weighted, controlled frames). In fact, the method extends to the n-sphere [3].

(2) *The spherical DWT*

There are many methods in the literature on approximation theory for designing a DWT on the two-sphere. Another one consists in lifting the plane 2-D DWT from the tangent plane at North Pole onto $\mathbb{S}^2$ using again the inverse *stereographic projection* [22]. This approach then yields locally supported orthonormal wavelet bases.

5.3 The CWT on other manifolds

The techniques developed for the two-sphere may be generalized to other manifolds, in particular, the so-called *conic sections* (the sphere, the two-sheeted hyperboloid, the paraboloid) [6] . In all cases, a key ingredient four designing a dilation *on* the manifold is a suitable projection on a simpler one (a tangent cone, a tangent plane) [7].

(1) *The two-sheeted hyperboloid*

Since the two-sheeted hyperboloid $\mathbb{H}^2$ is the dual of the two-sphere (in the sense of differential geometry), it is not surprising that a group-theoretical construction similar to the one on $\mathbb{S}^2$ works well, using isometry group SO(2,1) instead of SO(3). Thus one gets a fully developed CWT, but no results about frames are known [9].

(2) *The paraboloid*

The paraboloid $\mathbb{P}^2$ is a singular case, which can be obtained from both $\mathbb{S}^2$ and $\mathbb{H}^2$ by a limiting procedure. However, no (large) isometry group is available, so that the previous method do not work. An alternative construction is possible by lifting the CWT (or the DWT) vertically from the tangent plane onto the paraboloid [7].

(3) *The two-torus*

The two-torus $\mathbb{T}^2$ is isomorphic to the product of two circles, $\mathbb{T}^2 \simeq \mathbb{S}^1 \times \mathbb{S}^1 \simeq \mathrm{SO}(2) \times \mathrm{SO}(2)$. This fact suggests to exploit the WT on circle, which is analogous to that on the two-sphere, by the same construction, replacing SO(3,1) by SO(2,1). Now, since SO(2,2) $\simeq$ SO(2,1) $\times$ SO(2,1)/$\mathbb{Z}_2$, one may go one step further and use SO(2,2). However, this will introduce *two* independent dilations instead of a single, global one. This difficulty may then circumvented by combining SO(2,2) with the modular group $\mathrm{SL}(2, \mathbb{Z})$. In that way one may obtain both a CWT and discrete wavelet frames on $\mathbb{T}^2$ [12].

(4) *More general manifolds :* In the case of a general (two-dimensional) manifold $\mathcal{M}$, a possible approach is to lift *locally* the plane CWT from the tangent plane along thelocal normal, thus defining a genuine notion of local dilation on $\mathcal{M}$. In this way, one obtains a local CWT on $\mathcal{M}$. It remains to glue the local patches by standard methods of differential geometry, but is not necessarily easy [7].

6 Spatio-temporal wavelets and motion estimation

6.1 Introduction

There exist many methods for motion estimation : Optical flow, Block matching, Phase difference, ... An alternative approach is provided by the motion-tuned continuous wavelet transform, developed by M. Duval-Destin (1991), R.Murenzi (1992) and F. Mujica (1999). The idea is to adapt to the (2D+T) space-time the general formalism of the continuous wavelet transform on a manifold, described in Section 5.1 above. The rest of this section is based on our paper [11].

Technically speaking, one builds a time dependent wavelet, separable in frequency space:

$$\widehat{\psi}_{ST}(\vec{k}, \omega) = \underbrace{\widehat{\psi}_S(k_x, k_y)}_{\text{2-D wavelet}} \cdot \underbrace{\widehat{\psi}_T(\omega)}_{\text{1-D wavelet}} \quad . \tag{6.1}$$

Then one acts on it by the space-time (2D+T) group, containing space and time translations, space and time dilations, space rotations , namely, $G = \mathrm{SIM}(2) \times G_{\rm aff}^{+}$, via the usual unitary irreducible, square integrable, representations in the space of signals (image sequences) $L^2(\mathbb{R}^2 \times \mathbb{R}, d\vec{x}\, dt)$ described in Section 4.2. Finally, one replaces the separate space and time dilations a_s, a_t by a *global* dilation a and a *speed tuning* parameter c. With the Fourier transform written as

$$\widehat{s}(\vec{k}, \omega) = (2\pi)^{-3/2} \iint_{\mathbb{R}^2 \times \mathbb{R}} e^{-i(\vec{k}\cdot\vec{x} + \omega t)} s(\vec{x}, t)\, d\vec{x}\, dt \,,$$

the Fourier space is simply $L^2(\mathbb{R}^2 \times \mathbb{R}, d\vec{k}\, d\omega)$.

The principle underlying motion estimation is that speed detection and quantization is done in the Fourier space, because the wavelet measures the *inclination* of the spectrum. Indeed the

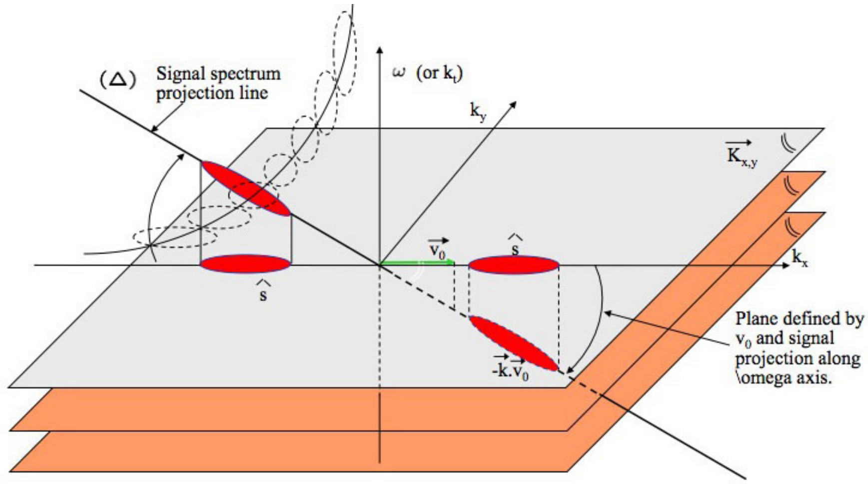


Figure 17: Inclination of a signal spectrum as a function of speed

inclination of a signal spectrum increases with speed, as shown in Fig. 17: a static object $\widehat{s}$ lives in the plane $\omega = 0$ of zero frequency, whereas an object $\widehat{s}$ moving with constant speed $\vec{v}$ lives in the plane $\vec{k} \cdot \vec{v} + \omega = 0$.

The next step is to identify the motion operators acting in Fourier space, namely,

$$\begin{array}{lrcl}
\text{Dilation :} & [\widehat{D}^a\, \widehat{\psi}](\vec{k}, \omega) & = & a^{3/2}\widehat{\psi}(a\vec{k}, a\,\omega) \\
\text{Translation :} & [\widehat{T}^{\vec{b},\tau}\, \widehat{\psi}](\vec{k}, \omega) & = & e^{-i(\vec{k}\cdot\vec{b}+\omega\tau)}\widehat{\psi}(\vec{k}, \omega) \\
\text{Rotation :} & [\widehat{R}^{\theta}\, \widehat{\psi}](\vec{k}, \omega) & = & \widehat{\psi}(r_{-\theta}\vec{k}, \omega) \\
\text{Speed tuning :} & [\widehat{\Lambda}^{c}\, \widehat{\psi}](\vec{k}, \omega) & = & \widehat{\psi}(c^{q}\vec{k}, c^{-p}\omega)
\end{array}$$

More precisely, these are unitary operators acting in the Fourier space $L^2(\mathbb{R}^2 \times \mathbb{R}, d\vec{k}\, d\omega)$.

In the definition of speed tuning, we have the additional constraints that $\widehat{\Lambda}^c$ must map the $\vec{v}_o$-plane, $\vec{k} \cdot \vec{v}_o + \omega = 0$, into the $c\vec{v}_o$-plane and must be unitary, which implies that $p = 2/3$ and

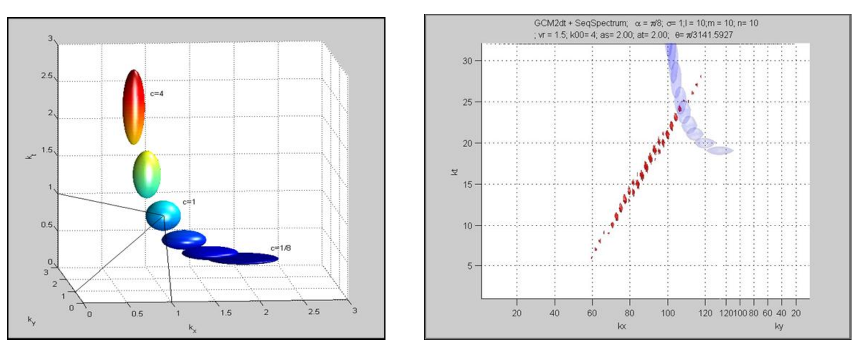


Figure 18: (Left)The hyperbola of speed-tuned wavelets; (Right) Speed analysis of an object moving at constant speed.

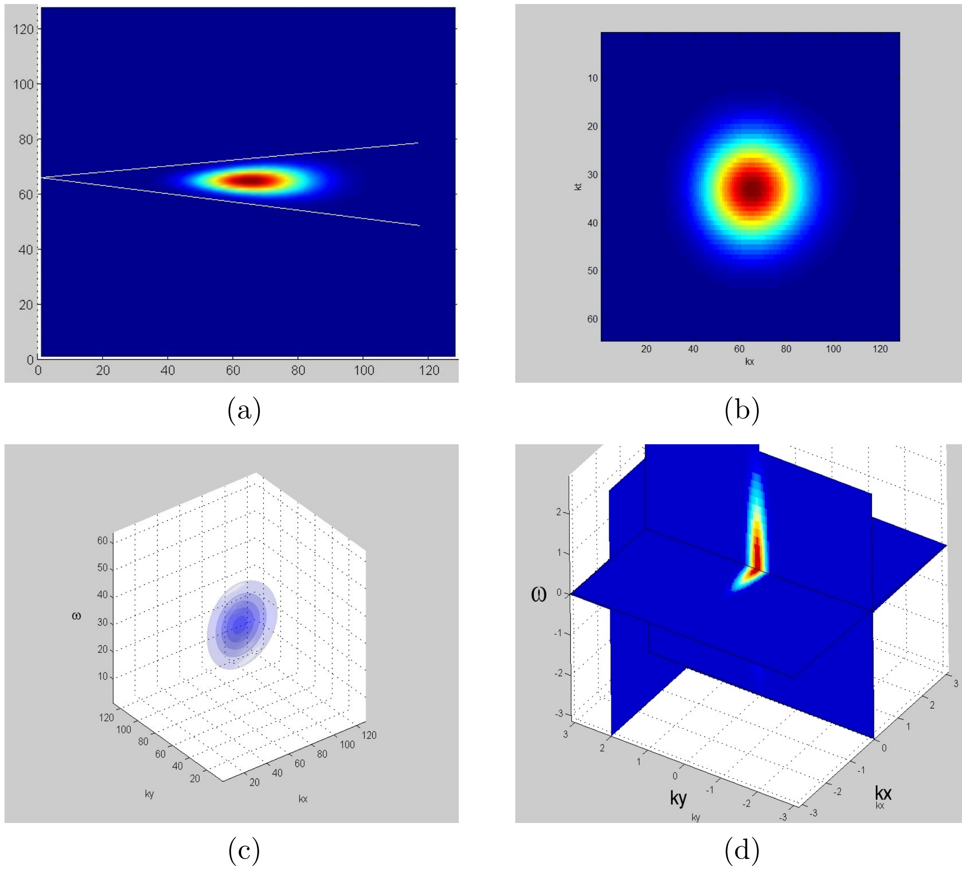


Figure 19: The GCM wavelet: (a) Top view ($c = 1, \alpha = \pi/16$); (b) Lateral view in plane (k_x, ω); (c) 3-D view of GCM in (k_x, k_y, ω) with $c = 1$; (d) 3-D sectional views in (k_x, k_y, ω) with $c = 2$)

$q = 1/3$. This correspond to the psycho-visual effect. If an object moves fast, only *large* details can be detected, but, if it moves slowly, then *small* details can be detected. Hence, a high speed object must be large to be "captured" and a low speed object can be small. Therefore wavelets must be speed-tuned (distorted and elongated) to "capture" a moving objet, which means that wavelets move on a hyperbola-like curve with increasing speed, as seen on the left panel of Fig. 18. On the right panel of the same figure, we show the speed analysis of an object moving at constant speed: capture is achieved when the signal spectrum (red) intersects the family of speed-tuned wavelets (blue).

6.2 Motion estimation with the GCM wavelet

The next step is to choose adequate wavelets for space and time components. The standard 2D+T wavelet is the Duval-Destin-Murenzi (DDM) wavelet

$$\widehat{\psi}_{DDM}(\vec{k}, \omega) = \underbrace{\widehat{\psi}_M(k_x, k_y)}_{\text{2-D Morlet wavelet}} \cdot \underbrace{\widehat{\psi}_M(\omega)}_{\text{1-D Morlet}}$$

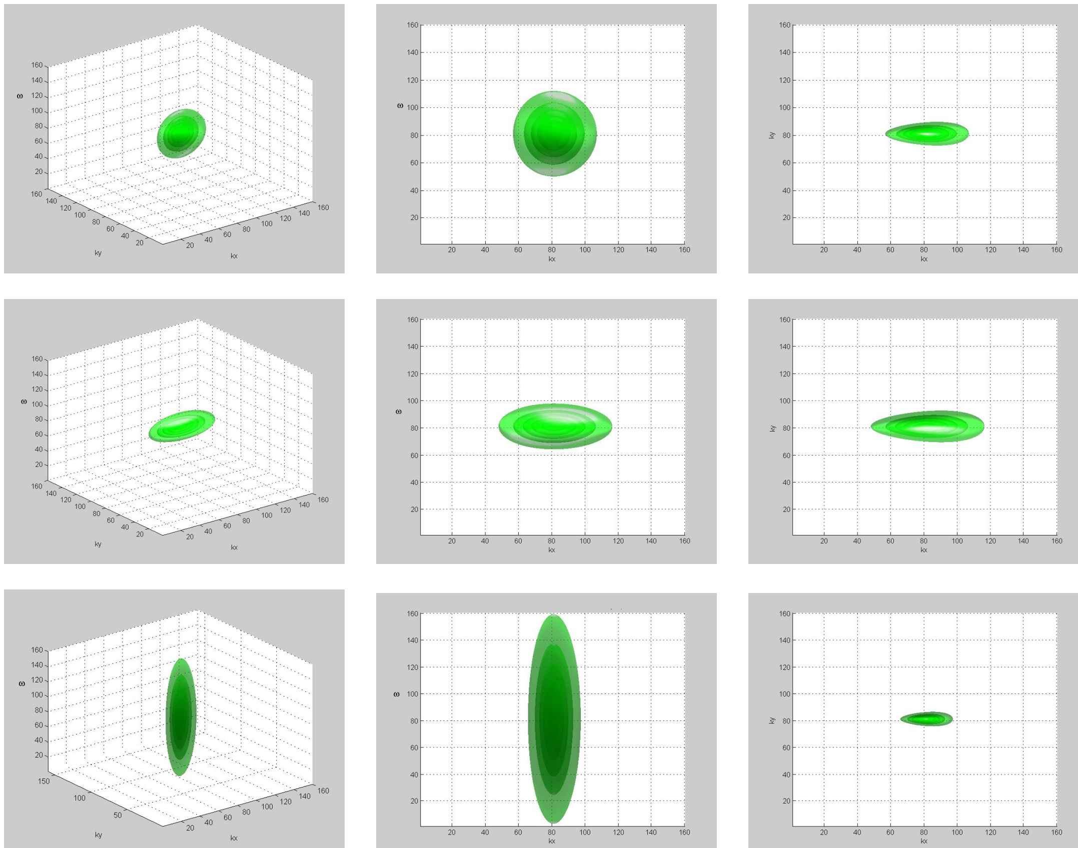


Figure 20: The GCM wavelet tuned to different velocities c : (Top row) $c = 1$; (Middle row) $c = 0.4$; (Bottom row) $c = 4$. The left column is a 3-D view, the central column is a side view in the (k_x, ω) plane, and the right column is a top view in the (k_x, k_y) plane.

where the 1-D Morlet wavelet is

$$\widehat{\psi}_M(\omega) = \exp(-\tfrac{1}{2}(\omega - \omega_0)^2) - \widehat{h}(\omega),$$

with the correction term $\widehat{h}(\omega)$ negligible in practice (for $\omega_0 \gtrsim 5.5$).

Since the 2-D Morlet wavelet has poor selectivity properties, we replace it by a *Gaussian-conical wavelet*, as given in (4.17), and get a GCM 2D+T wavelet

$$\widehat{\psi}_{lm}^{\mathrm{GCM}}(\vec{k}, \omega) = \begin{cases} \underbrace{\widehat{\psi}_{lm}^{GC}(k_x, k_y)}_{\text{2-D Gaussian-Conical}} \cdot \underbrace{\widehat{\psi}^{M}(\omega)}_{\text{1-D Morlet}}, \; \vec{k} \in \mathcal{C}(-\alpha, \alpha), \\ 0, \text{ otherwise,} \end{cases}$$

Explicitly

$$\widehat{\psi}_{lm}^{\mathrm{GCM}}(\vec{k}, \omega) = \begin{cases} (\vec{k} \cdot \vec{e}_{-\widetilde{\alpha}})^l \, (\vec{k} \cdot \vec{e}_{\widetilde{\alpha}})^m e^{-\frac{\sigma}{2}(k_x - \chi(\sigma))^2} \, e^{-\frac{1}{2}(\omega - \omega_0)^2}, \; \vec{k} \in \mathcal{C}(-\alpha, \alpha), \\ 0, \; \mathsf{otherwise} \end{cases}$$

Here the parameter $\sigma > 0$ controls the width of the Gaussian , $(\sqrt{l+m}, 0)$ is the central frequency and the *center correction term* $\chi(\sigma) = \sqrt{l+m}\frac{\sigma-1}{\sigma}$ controls the radial support of the wavelet $\widehat{\psi}_{lm}^{\rm GCM}$.

We show in Fig.19 the GCM wavelet in Fourier space. the bottom panels are 3-D views; in particular, the panel (d) shows the Gaussian behavior of the (vertical) Morlet part and the (horizontal) conical part.

In order to visualize the (hyperbolic) deformation of the wavelet, we show it tuned to different velocities c in Fig. 20.

The speed capture process runs as follows In order to capture the initial speed, independently of the object scale, it is useful to center the wavelet in the Fourier plane i.e., translating the wave-vector $\vec{k}$ by its central frequency $\vec{k}_0$ and canceling the term ω_0 in the Morlet part

$$\vec{k}_0 = \frac{1}{a^2}\frac{\sqrt{l+m}}{c^q}(\cos\theta, \sin\theta)$$

However, the modified GCM is now a simple filter (no oscillation), not a wavelet anymore ! The angular resolving power are, respectively, for the Morlet wavelet with for $k_0 \gg 1, ARP(\psi^M) = 2\cot^{-1}(k_0\sqrt{\epsilon})$ and for the GCM wavelet $ARP = 2\alpha$, the opening angle of the cone. As for frame bounds, see (2.17), estimates have been given by Murenzi for the 2-D Morlet wavelet, by Mujica for the DDM wavelet, and these are valid for GCM as well (but it is rather complicated!).

In order to proceed with the analysis algorithm we have to *discretize* the CWT, as follows.

- Given a sequence of N frames, corresponding to the time variables $\tau_i, i = 1, \ldots, N$, compute its discretized CWT for discrete speeds $c = c_j$: $\mathcal{W}_\psi(\vec{b}, \tau_i, \theta; a, c_j)$;
- Compute the energy density of the $i^{\rm th}$ frame $|\mathcal{W}_\psi(c_j, \tau_i)|^2$, taken as a function of speed c_j only;
- For a group of N_0 frames among the N frames of the sequence, compute the *total energy*

$$E_{\rm tot}(c_j) = \sum_{i\in N_0}\sum_{\vec{b}_{nm}} |\mathcal{W}_\psi(\vec{b}_{nm}, c_j, \tau_i)|^2, \tag{6.2}$$

  where $\{\vec{b}_{nm}\}$ is a discretized version of the $\vec{b}$-plane;
- Study the curve $f(c_j) := E_{\rm tot}(c_j)$: the maximum v_m is reached when the speed of the tuned wavelet c_j matches the real speed v_r of the object (see an example of a traveling 2-D Gaussian in Fig **??**).

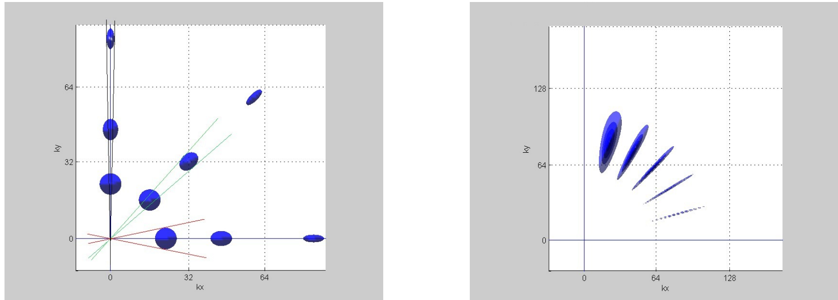
6.3 Experimentation

In this last section, we shall describe a number of experiments showing the excellent performance of the GCM wavelet for moton estimation.

(1) *Comparative aperture adjustments : Morlet vs. GCM*

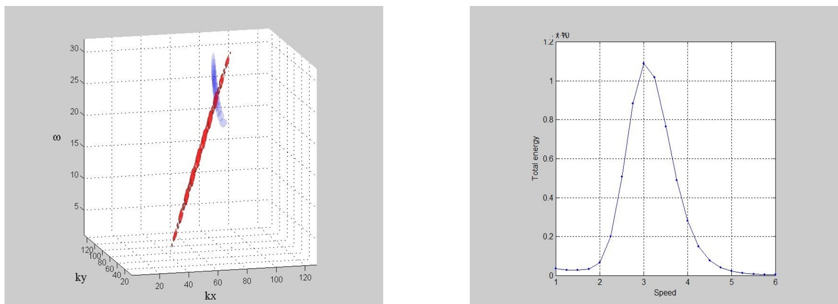
The first thing to check is for which wavelet, Morlet or GCM, the aperture is easier to adjust.

For small k_0, ϵ, the Morlet wavelet has very large aperture; the aperture decreases by increasing k_0, ϵ, but as k_0 increases, the Morlet wavelet moves away from the Fourier center (0,0), along its radius, which makes it very difficult to adjust the spatial positioning of this wavelet withe respect to a change in aperture selectivity. We show the effect in Fig. 21(a), on 3 different orientations, for the sake of clarity.



(a) (b)

Figure 21: Aperture adjustment for (a) the Morlet wavelet; and (b) the GCM wavelet



(a) (b)

Figure 22: (a) The spectrum of the sequence "traveling Gaussian" (red) intersects the hyperbola-like family of speed-tuned GCM wavelets (blue); (b) The total energy is maximum for $c = v_r = 3$.

For the GCM wavelet, on the contrary, the couple orientation/aperture is extremely simple to adjust. Fig. 21(b) shows 5 orientations $\pi/12 \leqslant \theta_m \leqslant 5\pi/12\}$ with the GCM tuned to aperture in the range $\pi/256 \leqslant \alpha_m \leqslant \pi/16$: the wavelet does not move, it is simply squeezed.

The conclusion, obviously, is that the GCM wavelet is superior to the 2D+T Morlet wavelet for object recognition and tracking by spectral signature.

(2) *Velocity capture comparison*

In order to test the capability of our two wavelets for measuring the velocity of a moving object, we use a test sequence of $128 \times 128 \times 32$ describing the motion, at constant speed $v_r =$ 3 pixels/fr, of a 2-D Gaussian; the angles of the Gaussian and of its trajectory are varied (along OX, at 45° and along OY). It turns out that speed capture equally good with GCM and with Morlet in the case of an isotropic Gaussian. But, for an anisotropic Gaussian, with large σ_y and small σ_x (i.e. the spectrum is narrow along k_x), the test shows that the GCM wavelet has a far better aperture selectivity, i.e. its accuracy in directional speed capture is much better. Fig. 22(a) shows that the spectrum of the sequence "traveling Gaussian" (red) intersects the hyperbola-like family of speed-tuned GCM wavelets (blue). Panel (b) then shows the total energy (6.2) is indeed maximum for $c = v_r = 3$.

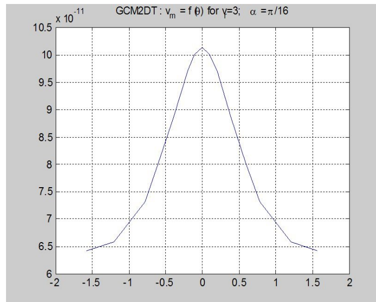


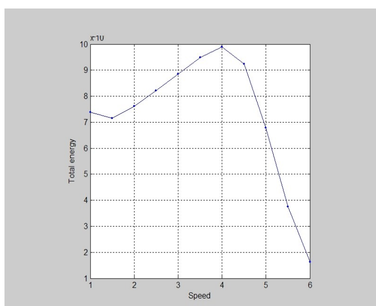
Figure 23: Curve v_m vs. θ_{wav} for $-\pi/2 \leqslant \theta \leqslant +\pi/2$, with GCM.

(3) *Orientation accuracy in directional speed capture*

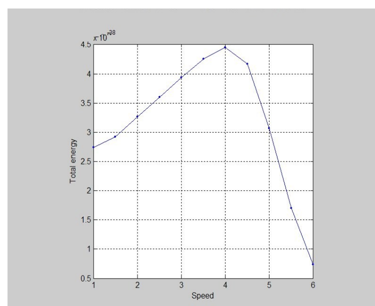
In order to test the performance of the GCM wavelet for measuring the direction of the mobile velocity, we show in Fig. 23 the curve v_m vs. θ_{wav} for $-\pi/2 \leqslant \theta \leqslant +\pi/2$. The outcome is that the correct speed is captured when the wavelet orientation ($\theta = 0$) exactly corresponds to the spectrum orientation (along k_x). Thus, contrary to Morlet, GCM has a good angular selectivity and is efficient at detecting the correct speed of the sequence in a very narrow angular aperture and not for other directions.

(4) *Aperture accuracy in speed capture*

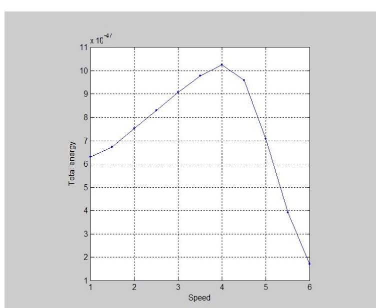
In the next experiment (Fig. 24), we test again the speed capture performance of the GCM



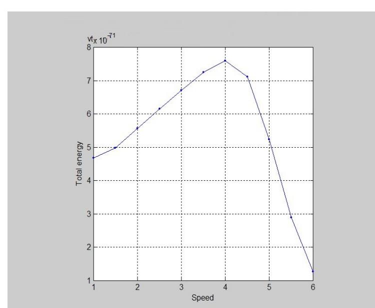
(a)



(b)



(c)



(d)

Figure 24: Speed detection with the GCM wavelet for increasingly small aperture of the cone: (a) $\pi/8$; (b) $\pi/16$; (c) $\pi/64$; and (d) $\pi/256$.

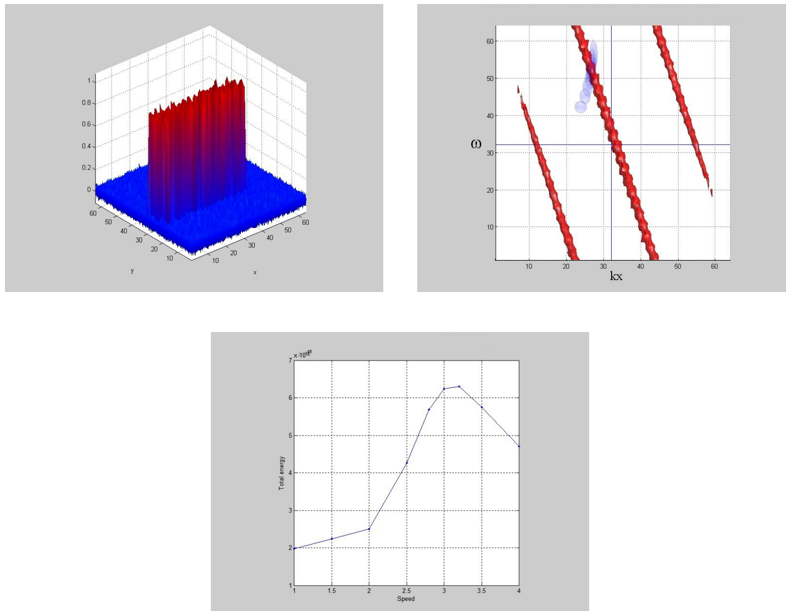


Figure 25: Speed detection in the presence of noise: (Top) The noisy sequence and the intersection of the sequence spectrum (red) with the hyperbola of GCM speed-tuned wavelets (blue); (Bottom) The energy as a function of c.

wavelet, at $v_r = 4$ and for increasingly narrow conical apertures : $\pi/8, \pi/16, \pi/64, \pi/256\,(0.70°)$. The result is that GCM captures the spectrum for extremely weak apertures ($\pi/256 = (0.70°!)$ without moving in the Fourier space. Thus, contrary to Morlet, GCM is very robust to scale initialization and object size detection and tracking.

(5) *Stability with respect to noise*

In our last experiment, we test the stability of the GCM wavelet with respect to noise (Fig. 25). To that effect, we choose for test sequence a rectangle of size 9×2 pixels in horizontal translation at a speed of 3 pixels/fr, perturbed by white Gaussian noise. The noise level ranges from an SNR of 50 dB up to 25 dB. The outcome is that the correct speed captured up to a noise level of 32 dB, in the sense the energy curve has its maximum precisely at $c = 3$.

6.4 Conclusion

In conclusion, we may say that the GCM wavelet is a highly directionally selective speed-tuned wavelet, which provides a much more powerful tool than the 2D+T Morlet wavelet for the recognition and tracking of spectral signatures. Indeed, it shows an extreme efficiency in directional speed selectivity down to angle apertures of less than 1 degree ($\pi/256$) and a good capacity of radial stability and adjustment with respect to variation of the aperture.

Future work along this line may envisage target tracking with GCM instead of DDM, speed detection/quantization in video sequences, explicit or estimated frame bounds for GCM, trajectory detection in direct space with GCM, ... Another possible extension of this work is to exploit curvelets and shearlets for motion analysis, but this will probably be much more complex.

References

[1] S.T. Ali, J-P. Antoine and J-P. Gazeau, *Coherent States, Wavelets and Their Generalizations*, 2nd ed. (Springer-Verlag, New York et al., 2013). 1, 2.4, 2.5

[2] J-P. Antoine, R. Murenzi, P. Vandergheynst and S.T. Ali, *Two-dimensional Wavelets and Their Relatives* (Cambridge University Press, Cambridge (UK), 2004). 1, 4.6

[3] J-P. Antoine and P. Vandergheynst, Wavelets on the n-sphere and related manifolds, *J. Math. Phys.* **39** (1998) 3987–4008. 5.2

[4] J-P. Antoine and P. Vandergheynst, Wavelets on the 2-sphere: A group-theoretical approach, *Appl. Comput. Harmon. Anal.* **7** (1999) 262–291. 5.2

[5] J-P. Antoine, L. Demanet, L. Jacques, and P. Vandergheynst, Wavelets on the sphere: Implementation and approximations, *Appl. Comput. Harmon. Anal.* **13** (2002) 177–200 . 5.2

[6] J-P. Antoine, I. Bogdanova, and P. Vandergheynst, The continuous wavelet transform on conic sections, *Int. J. Wavelets, Multires. and Inform. Proc.*, **6** (2007) 137–156. 5.3

[7] J-P. Antoine, D. Roşca, and P. Vandergheynst, Wavelet transform on manifolds: Old and new approaches, *Appl. Comput. Harmon. Anal.* **28** (2010) 189–202. 5.3

[8] I. Bogdanova, P. Vandergheynst, J.-P. Antoine, L. Jacques, and M. Morvidone, Stereographic wavelet frames on the sphere, *Appl. Comput. Harmon. Anal.* **19** (2005) 223–252. 5.2

[9] I. Bogdanova, P. Vandergheynst, and J-P. Gazeau, Continuous wavelet transform on the hyperboloid, *Appl. Comput. Harmon. Anal.* **23** (2007) 286–306. 5.3

[10] J. C. van den Berg (ed.), *Wavelets in Physics* (Cambridge Univ. Press, Cambridge (UK), 1999). 1

[11] P. Brault and J-P. Antoine, A spatio-temporal Gaussian-Conical wavelet with high aperture selectivity for motion and speed analysis, *Applied Comput. Harmon. Anal.* **34** (2013) 148–161. 6.1

[12] M. Calixto, J. Guerrero, and D. Roşca, Wavelet transform on the torus: a group-theoretical approach, preprint (2013). 5.3

[13] I. Daubechies, *Ten Lectures on Wavelets* (SIAM, Philadelpia, 1992). 1, 2.6

[14] D. Gabor, Theory of communication, *J. Inst. Electr. Engrg.(London)* **93** (1946) 429–457. 1

[15] C. Heil and D. Walnut, Continuous and discrete wavelet transforms, *SIAM Review* **31** (1989) 628–666. 1

[16] B. B. Hubbard, *The World According to Wavelets*, 2nd ed. (A.K. Peters, Wellesley, MA, 1998). 1

[17] C. Lemke, A. Schuck Jr., J-P. Antoine, and D. Sima, Metabolite-sensitive analysis of magnetic resonance spectroscopic signals using the continuous wavelet transform, *Meas. Sci. Technol.* **22** (2011) 114013. 2.5

[18] S. G. Mallat, *A Wavelet Tour of Signal Processing, 3nd ed. : The Sparse Way* (Academic Press, San Diego, 2008). 1

[19] D. Marr, *Vision* (Freeman, San Francisco, 1982). 4.5

[20] Y. Meyer, *Wavelets: Algorithms and Applications* (SIAM, Philadelphia, PA, 1993). 1

[21] O. Rioul and M. Vetterli, Wavelets and signal processing, *IEEE SP Magazine,* October 1991, 14–38. 1

[22] D. Roşca and J-P. Antoine, Locally supported orthogonal wavelet bases on the sphere via stereographic projection, *Math. Probl. Engineering,* Vol. 2009 (2009) 124904 5.2

[23] A. Suvichakorn, H. Ratiney, A. Bucur, S. Cavassila, and J-P. Antoine, Toward a quantitative analysis of *in vivo* magnetic resonance proton spectroscopic signals using the continuous Morlet wavelet transform, *Meas. Sci. Technol.* **20** (2009) 104029. 2.5

[24] A. Suvichakorn, C. Lemke, A. Schuck Jr. and J-P. Antoine, The Continuous Wavelet Transform in MRS, Tutorial text, *Marie Curie Research Training Network FAST* (2011) `http://www.fast-mariecurie-rtn-project.eu/#Wavelet`. 2.5

[25] B. Torrésani, *Analyse continue par ondelettes* (InterÉditions/CNRS Éditions, Paris, 1995). 1

[26] J. Ville, Théorie et applications de la notion de signal analytique, *Câbles et Transm.* **2ème A** (1948) 61–74 . 1

[27] E. P. Wigner, On the quantum correction for thermodynamic equilibrium, *Phys. Rev.* **40**, 749–759 (1932). 1

[28] W. Wisnoe, Utilisation de la méthode de transformée en ondelettes 2D pour l'analyse de visualisation d'écoulements. Thèse de Doctorat ENSAE, Toulouse (1993). 4.6, 15, 16

[29] Wavelet Toolbox YAWTb: `http://sites.uclouvain.be/ispgroup/yawtb/`. 1