Wavelets and Signal Processing

OLIVIER RIOUL and MARTIN VETTERLI

Wavelet theory provides a unified framework for a number of techniques which had been developed independently for various signal processing applications. For example, multiresolution signal processing, used in computer vision; subband coding, developed for speech and image compression; and wavelet series expansions, developed in applied mathematics, have been recently recognized as different views of a single theory.

In fact, wavelet theory covers quite a large area. It treats both the continuous and the discrete-time cases. It provides very general techniques that can be applied to many tasks in signal processing, and therefore has numerous potential applications.

In particular, the Wavelet Transform (WT) is of interest for the analysis of *non-stationary* signals, because it provides an alternative to the classical Short-Time Fourier Transform (STFT) or Gabor transform [GAB46, ALL77, POR80]. The basic difference is as follows. In contrast to the STFT, which uses a single analysis window, the WT uses short windows at high frequencies and long windows at low frequencies. This is in the spirit of so-called "constant-Q" or constant relative bandwidth frequency analysis. The WT is also related to time-frequency analysis based on the Wigner-Ville distribution [FLA89, FLA90, RIO90a].

For some applications it is desirable to see the WT as a signal decomposition onto a set of basis functions. In fact, basis functions called *wavelets* always underlie the wavelet analysis. They are obtained from a single prototype wavelet by dilations and contractions (scal-

1053-5888/91/1000-0014$1.00©IEEE

ings) as well as shifts. The prototype wavelet can be thought of as a bandpass filter, and the constant-Q property of the other bandpass filters (wavelets) follows because they are scaled versions of the prototype.

Therefore, in a WT, the notion of *scale* is introduced as an alternative to frequency, leading to a so-called *time-scale representation*. This means that a signal is mapped into a time-scale plane (the equivalent of the time-frequency plane used in the STFT).

There are several types of wavelet transforms, and, depending on the application, one may be preferred to the others. For a continuous input signal, the time and scale parameters can be continuous [GRO89], leading to the Continuous Wavelet Transform (CWT). They may as well be discrete [DAU88, MAL89b, MEY89, DAU90a], leading to a Wavelet Series expansion. Finally, the wavelet transform can be defined for discrete-time signals [DAU88, RIO90b, VET90b], leading to a Discrete Wavelet Transform (DWT). In the latter case it uses multirate signal processing techniques [CRO83] and is related to subband coding schemes used in speech and image compression. Notice the analogy with the (Continuous) Fourier Transform, Fourier Series, and the Discrete Fourier Transform.

Wavelet theory has been developed as a unifying framework only recently, although similar ideas and constructions took place as early as the beginning of the century [HAA10, FRA28, LIT37, CAL64]. The idea of looking at a signal at various scales and analyzing it with various resolutions has in fact emerged independently in many different fields of mathematics, physics and engineering. In the mid-eighties, researchers of the "French school," lead by a geophysicist, a theoretical physicist and a mathematician (namely, Morlet, Grossmann, and Meyer), built strong mathematical foundations around the subject and named their work "*Ondelettes*" (Wavelets). They also interacted considerably with other fields.

The attention of the signal processing community was soon caught when Daubechies and Mallat, in addition to their contribution to the theory of wavelets, established connections to discrete signal processing results [DAU88], [MAL89a]. Since then, a number of theoretical, as well as practical contributions have been made on various aspects of WT's, and the subject is growing rapidly [WAV89], [IT92].

The present paper is meant both as a review and as a tutorial. It covers the main definitions and properties of wavelet transforms, shows connections among the various fields where results have been developed, and focuses on signal processing applications. Its purpose is to present a simple, synthetic view of wavelet theory, with an easy-to-read, non-rigorous flavor. An extensive bibliography is provided for the reader who wants to go into more detail on a particular subject.

NON-STATIONARY SIGNAL ANALYSIS

The aim of signal analysis is to extract relevant information from a signal by transforming it. Some methods make *a priori* assumptions on the signal to be analyzed; this may yield sharp results if these assumptions are valid, but is obviously not of general applicability. In this paper we focus on methods that are applicable to any general signal. In addition, we consider invertible transformations. The analysis thus unambiguously represents the signal, and more involved operations such as parameter estimation, coding and pattern recognition can be performed on the "transform side," where relevant properties may be more evident.

Such transforms have been applied to *stationary* signals, that is, signals whose properties do not evolve in time (the notion of stationarity is formalized precisely in the statistical signal processing literature). For such signals $x(t)$, the natural "stationary transform" is the well-known Fourier transform [FOU88]:

$$X(f) = \int_{-\infty}^{+\infty} x(t)\, e^{-2j\pi f t} dt \qquad (1)$$

The analysis coefficients $X(f)$ define the notion of global frequency f in a signal. As shown in (1), they are computed as inner products of the signal with sinewave basis functions of infinite duration. As a result, Fourier analysis works well if $x(t)$ is composed of a few stationary components (e.g., sinewaves). However, any abrupt change in time in a non-stationary signal $x(t)$ is spread out over the whole frequency axis in $X(f)$. Therefore, an analysis adapted to *nonstationary* signals requires more than the Fourier Transform.

The usual approach is to introduce time dependency in the Fourier analysis while preserving linearity. The idea is to introduce a "local frequency" parameter (local in time) so that the "local" Fourier Transform looks at the signal through a window over which the signal is approximately stationary. Another, equivalent way is to modify the sinewave basis functions used in the Fourier Transform to basis functions which are more concentrated in time (but less concentrated in frequency).

SCALE VERSUS FREQUENCY

The Short-Time Fourier Transform: Analysis with Fixed Resolution.

The "instantaneous frequency" [FLA89] has often been considered as a way to introduce frequency de-

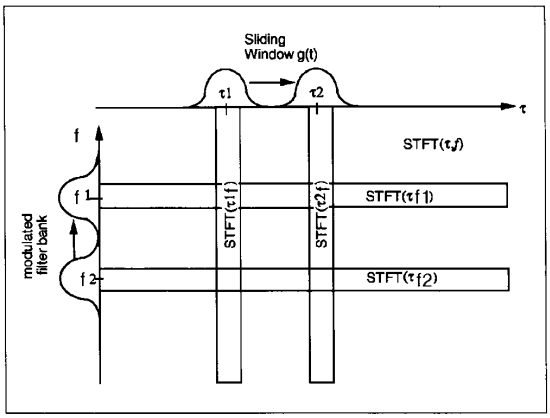


Fig. 1. Time-frequency plane corresponding to the Short-Time Fourier Transform. It can be seen either as a succession of Fourier Transforms of a windowed segment of the signal (vertical stripes) or as a modulated analysis filter bank (horizontal stripes).

pendence on time. If the signal is not narrow-band, however, the instantaneous frequency averages different spectral components in time. To become accurate in time, we therefore need a *two-dimensional* time-frequency representation $S(t,f)$ of the signal $x(t)$ composed of spectral characteristics depending on time, the local frequency f being defined through an appropriate definition of $S(t,f)$. Such a representation is similar to the notation used in a musical score, which also shows "frequencies" played in time.

The Fourier Transform (1) was first adapted by Gabor [GAB46] to define $S(t,f)$ as follows. Consider a signal $x(t)$, and assume it is stationary when seen through a window $g(t)$ of limited extent, centered at time location τ. The Fourier Transform (1) of the windowed signals $x(t)\, g^*(t-\tau)$ yields the Short-Time Fourier Transform (STFT)

$$\mathrm{STFT}(\tau, f) = \int x(t)\, g^*(t-\tau)\, e^{-2j\pi f t}\, dt \qquad (2)$$

which maps the signal into a two-dimensional function in a time-frequency plane (τ,f). Gabor originally only defined a synthesis formula, but the analysis given in (2) follows easily.

The parameter f in (2) is similar to the Fourier frequency and many properties of the Fourier transform carry over to the STFT. However, the analysis here depends critically on the choice of the window $g(t)$.

Figure 1 shows vertical stripes in the time-frequency plane, illustrating this "windowing of the signal" view of the STFT. Given a version of the signal windowed around time t, one computes all "frequencies" of the STFT.

An alternative view is based on a filter bank interpretation of the same process. At a given frequency f, (2) amounts to filtering the signal "at all times" with a bandpass filter having as impulse response the window function modulated to that frequency. This is shown as the horizontal stripes in Fig. 1. Thus, the STFT may be seen as a modulated filter bank [ALL77], [POR80].

From this dual interpretation, a possible drawback related to the time and frequency resolution can be shown. Consider the ability of the STFT to discriminate between two pure sinusoids. Given a window function $g(t)$ and its Fourier transform $G(f)$, define the "bandwidth" Δf of the filter as

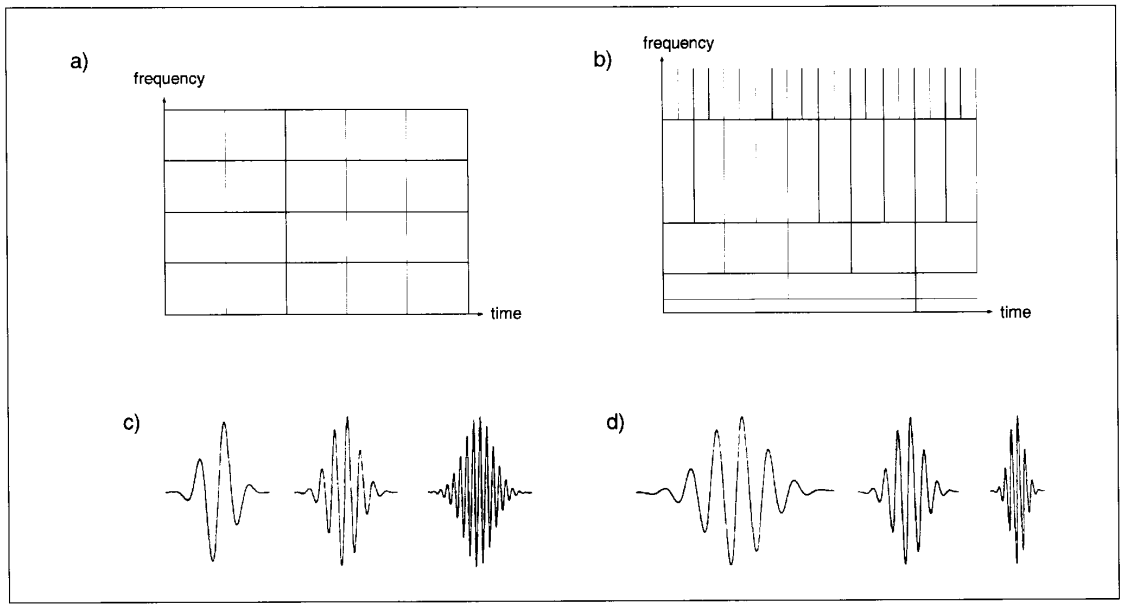


Fig. 2. Basis functions and time-frequency resolution of the Short-Time Fourier Transform (STFT) and the Wavelet Transform (WT). The tiles represent the essential concentration in the time-frequency plane of a given basis function. (a) Coverage of the time-frequency plane for the STFT, (b) for the WT. (c) Corresponding basis functions for the STFT, (d) for the WT ("wavelets").

$$\Delta f^2 = \frac{\int f^2 \, |G(f)|^2 \, df}{\int |G(f)|^2 \, df} \tag{3}$$

where the denominator is the energy of $g(t)$. Two sinusoids will be discriminated only if they are more than Δf apart (This is an rms measure, and others are possible). Thus, the resolution in frequency of the STFT analysis is given by Δf. Similarly, the spread in time is given by Δt as

$$\Delta t^2 = \frac{\int t^2 |g(t)|^2 \, dt}{\int |g(t)|^2 \, dt} \tag{4}$$

where the denominator is again the energy of $g(t)$. Two pulses in time can be discriminated only if they are more than Δt apart.

Now, resolution in time and frequency cannot be arbitrarily small, because their product is lower bounded.

$$\text{Time} - \text{Bandwidth product} = \Delta t \, \Delta f \geq \frac{1}{4\pi} \tag{5}$$

This is referred to as the uncertainty principle, or Heisenberg inequality. It means that one can only trade time resolution for frequency resolution, or *vice versa*. Gaussian windows are therefore often used since they meet the bound with equality [GAB46].

More important is that once a window has been chosen for the STFT, then the time-frequency resolution given by (3), (4) is *fixed* over the entire time-frequency plane (since the same window is used at all frequencies). This is shown in Fig. 2a, while Fig. 2c shows the associated basis functions of the STFT. For example, if the signal is composed of small bursts associated with long quasi-stationary components, then each type of component can be analyzed with good time resolution or frequency resolution, but not both.

The Continuous Wavelet Transform: A Multiresolution Analysis.

To overcome the resolution limitation of the STFT, one can imagine letting the resolution Δt and Δf vary in the time-frequency plane in order to obtain a multiresolution analysis. Intuitively, when the analysis is viewed as a filter bank, the time resolution must increase with the central frequency of the analysis filters. We therefore impose that Δf is proportional to f, or

$$\frac{\Delta f}{f} = c \tag{6}$$

where c is a constant. The analysis filter bank is then composed of band-pass filters with constant relative bandwidth (so-called "constant-Q" analysis). Another way to say this is that instead of the frequency responses of the analysis filter being regularly spaced over the frequency axis (as for the STFT case), they are regularly spread in a logarithmic scale (see Fig. 3). This kind of filter bank is used, for example, for modeling the frequency response of the cochlea situated in the inner ear and is therefore adapted to auditory perception, e.g. of music: filters satisfying (6) are naturally distributed into octaves.

When (6) is satisfied, we see that Δf and therefore also Δt changes with the center frequency of the analysis filter. Of course, they still satisfy the Heisenberg inequality (5), but now, the time resolution becomes arbitrarily good at high frequencies, while the frequency resolution becomes arbitrarily good at low frequencies. For example, two very close short bursts can always be eventually separated in the analysis by going up to

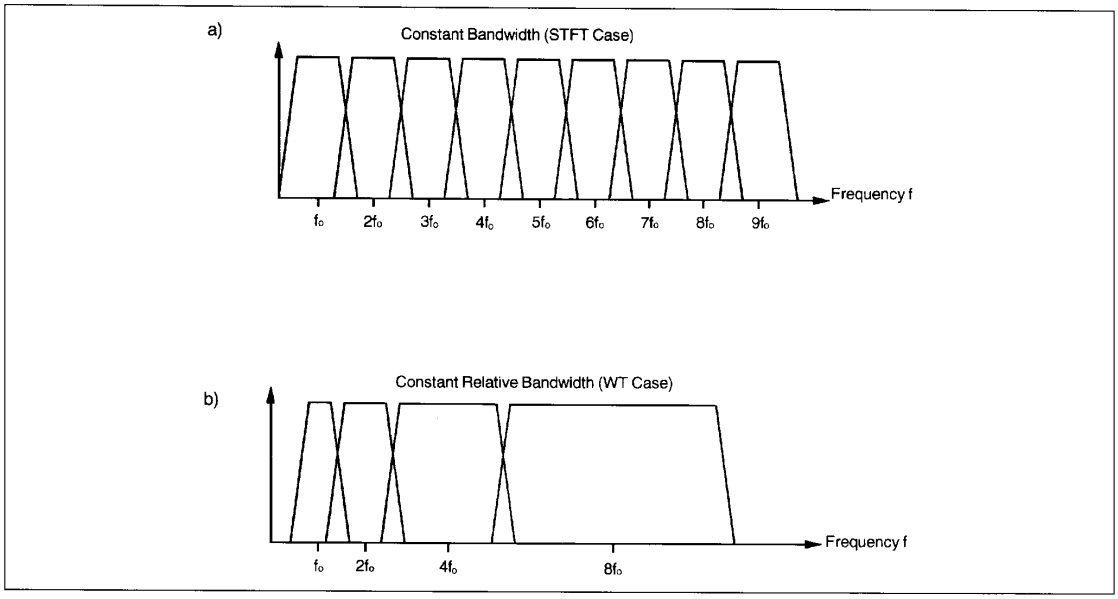


Fig. 3. Division of the frequency domain (a) for the STFT (uniform coverage) and (b) for the WT (logarithmic coverage).

Box 1: The Notion of Scale and Resolution

First, recall that when a function $f(t)$ is scaled:

$f(t) \rightarrow f(at)$, *where* $a > 0$,

then it is contracted if $a > 1$ and expanded if $a < 1$ Now, the CWT can be written either as

$$\mathrm{CWT}_x(\tau,a) = \frac{1}{\sqrt{a}} \int x(t)\, h^*(\frac{t-\tau}{a})dt \qquad \text{(B1.1)}$$

or, by a change of variable, as

$$\mathrm{CWT}_x(\tau,a) = \sqrt{a} \int x(at)\, h^*(t - \frac{\tau}{a})dt \qquad \text{(B1.2)}$$

The interpretation of (B1.1) is that as the scale increases, the filter impulse response $h\left(\frac{t-\tau}{a}\right)$ becomes spread out in time, and takes only long-time behavior into account. Equivalently, (B1.2) indicates that as the scale grows, an increasingly contracted version of the signal is seen through a constant length filter. That is, the scale factor a has the interpretation of the scale in maps. Very large scales mean global views, while very small scales mean detailed views.

A related but different notion is that of resolution. The resolution of a signal is linked to its frequency content. For example, lowpass filtering a signal keeps its scale, but reduces its resolution.

Scale change of continuous time signals does not alter their resolution, since the scale change can be reversed. However, in discrete-time signals, increasing the scale in the analysis involves subsampling, which automatically reduces the resolution. Decreasing the scale (which involves upsampling) can be undone, and does not change the resolution. The interplay of scale and resolution changes in discrete-time signals is illustrated in Fig. 9 and fully explained in [RIO90b], [VET90b].

higher analysis frequencies in order to increase time resolution (see Fig. 2b). This kind of analysis of course works best if the signal is composed of high frequency components of short duration plus low frequency components of long duration, which is often the case with signals encountered in practice.

A generalization of the concept of changing resolution at different frequencies is obtained with so-called "wavelet packets" [WIC89], where arbitrary time-frequency resolutions (within the uncertainty bound (5)) are chosen depending on the signal.

The *Continuous Wavelet Transform* (CWT) exactly follows the above ideas while adding a simplification: all impulse responses of the filter bank are defined as *scaled* (i.e. stretched or compressed) versions of the same prototype $h(t)$, i.e.,

$$h_a(t) = \frac{1}{\sqrt{|a|}}\, h(\frac{t}{a})$$

where a is a *scale factor* (the constant $1/\sqrt{|a|}$ is used for energy normalization). This results in the definition of the CWT:

$$\mathrm{CWT}_x(\tau,\, a) = \frac{1}{\sqrt{|a|}} \int x(t)\, h^*\left(\frac{t-\tau}{a}\right)dt \qquad (7)$$

Since the same prototype $h(t)$, called the *basic wavelet*, is used for all of the filter impulse responses, no specific scale is privileged, i.e. the wavelet analysis is self-similar at all scales. Moreover, this simplification is useful when deriving mathematical properties of the CWT.

To make the connection with the modulated window used in the STFT clearer, the basic wavelet $h(t)$ in (7) could be chosen as a modulated window [GOU84, GRO84, GRO89]

$$h(t) = g(t)\, e^{-2j\pi f_0 t}$$

Then the frequency responses of the analysis filters indeed satisfy (6) with the identification

$$a = \frac{f_0}{f}$$

But more generally, $h(t)$ can be any band-pass function and the scheme still works. In particular one can dispense with complex-valued transforms and deal only with real-valued ones.

It is important to note that here, the local frequency $f = a f_0$ has little to do with that described for the STFT: indeed, it is associated with the scaling scheme (see Box 1). As a result, this local frequency, whose definition depends on the basic wavelet, is no longer linked to frequency modulation (as was the case for the STFT) but is now related to time-scalings. This is the reason why the terminology "scale" is often preferred to "frequency" for the CWT, the word "frequency" being reserved for the STFT. Note that we define *scale* in wavelet analysis like the scale in geographical maps: since the filter bank impulse responses in (7) are dilated as scale increases, large scale corresponds to contracted signals, while small scale corresponds to dilated signals.

WAVELET ANALYSIS AND SYNTHESIS

Another way to introduce the CWT is to define *wavelets* as basis functions. In fact, basis functions already appear in the preceding definition (7) when one sees it as an inner product of the form

$$\mathrm{CWT}_x(\tau,\, a) = \int x(t)\, h^*_{a,\tau}(t)\, dt$$

which measures the "similarity" between the signal and the basis functions

Box 2: STFTs and CWTs as Cross-Ambiguity Functions

The inner product is often used as a similarity measurement, and because both STFT's and CWT's are inner products, they appear in several detection/estimation problems. Consider, for example, the problem of estimating the location and velocity of some target in radar or sonar applications. The estimation procedure consists in first emitting a known signal $h(t)$. In the presence of a target, this signal will return to the source (received signal $x(t)$) with a certain delay τ, due to the target's location, and a certain distortion (Doppler effect), due to the target's velocity.

For narrow-band signals, the Doppler effect amounts to a single frequency shift f_0 and the characteristics of the target will be determined by maximizing the cross-correlation function (called "narrow-band cross-ambiguity function") [WOO53]

$$\int x(t)\, h(t-\tau)\, e^{-2j\pi f_0 t}\, dt = \mathrm{STFT}(\tau, f)$$

For wide-band signals, however, the Doppler frequency shift varies in the signal's spectrum, causing a stretching or a compression in the signal. The estimator thus becomes the "wide-band cross-ambiguity function" [SPE67], [AUS90]

$$\frac{1}{\sqrt{|a|}} \int x(t)\, h\left(\frac{t-\tau}{a}\right) dt = \mathrm{CWT}_x(\tau, a)$$

As a result, in both cases, the "maximum likelihood" estimator takes the form of a STFT or a CWT, i.e. of an inner product between the received signal and either STFT or wavelet basis functions. The basis function which best fits the signal is used to estimate the parameters.

Note that, although the wide-band cross-ambiguity function is a CWT, for physical reasons, the dilation parameter a stays on the order of magnitude of 1, whereas it may cover several octaves when used in signal analysis [FLA89].

$$h_{a,\tau} = \frac{1}{\sqrt{a}}\, h\left(\frac{t-\tau}{a}\right)$$

called *wavelets*. The wavelets are scaled and translated versions of the basic wavelet prototype $h(t)$ (see Fig. 2d).

Of course, basis functions can be considered for the STFT as well. For both the STFT and the CWT, the sinewaves basis functions of the Fourier Transform are replaced by more localized reference signals labelled by time and frequency (or scale) parameters. In fact both transforms may be interpreted as special cases of the cross-ambiguity function used in radar or sonar processing (see Box 2).

The wavelet analysis results in a set of wavelet coefficients which indicate how close the signal is to a particular basis function. Thus, we expect that any general signal can be represented as a decomposition into wavelets, i.e. that the original waveform is synthesized by adding elementary building blocks, of constant shape but different size and amplitude. Another way to say this is that we want the continuously labelled wavelets $h_{a,\tau}(t)$ to behave just like an *orthogonal basis* [MEY90]. The analysis is done by computing inner products, and the synthesis consists of summing up all the orthogonal projections of the signal onto the wavelets.

$$x(t) = c \int_{a>0} \int \mathrm{CWT}(\tau, a)\, h_{a,\tau}(t)\, \frac{da\, d\tau}{a^2} \qquad (8)$$

where c is a constant that depends only on $h(t)$. The measure in this integration is formally equivalent to $dt\, df$ [GOU84]. We have assumed here that both signal and wavelets are either real-valued or complex analytic so that only positive dilations $a > 0$ have to be taken into account. Otherwise (8) is more complicated [GRO84].

Of course, the $h_{a,\tau}(t)$ are certainly not orthogonal since they are very redundant (they are defined for continuously varying a and τ). But surprisingly, the reconstruction formula (8) is indeed satisfied whenever $h(t)$ is of finite energy and *band pass* (which implies that it oscillates in time like a short wave, hence the name "wavelet"). More precisely, if $h(t)$ is assumed sufficiently regular, then the reconstruction condition is $\int h(t)\, dt = 0$.

Note that the reconstruction takes place only in the sense of the signal's energy. For example, a signal may be reconstructed only with zero mean since $\int h(t)\, dt = 0$. In fact the type of convergence of (8) may be strengthened and is related to the numerical robustness of the reconstruction [DAU90a].

Similar reconstruction can be considered for the STFT, and the similarity is remarkable [DAU90a]. However, in the STFT case, the reconstruction condition is less restrictive: only finite energy of the window is required.

SCALOGRAMS

The *spectrogram*, defined as the square modulus of the STFT, is a very common tool in signal analysis because it provides a distribution of the energy of the signal in the time-frequency plane. A similar distribution can be defined in the wavelet case. Since the CWT behaves like an orthonormal basis decomposition, it can be shown that it is isometric [GRO84], i.e., it preserves energy. We have

$$\int\int |\mathrm{CWT}(\tau, a)|^2 \frac{d\tau}{a^2}\, \frac{da}{a^2} = E_x$$

where $E_x = \int |x(t)|^2\, dt$ is the energy of the signal $x(t)$. This leads us to define the *wavelet spectrogram*, or *scalogram*, as the squared modulus of the CWT. It is a distribution of the energy of the signal in the time-scale

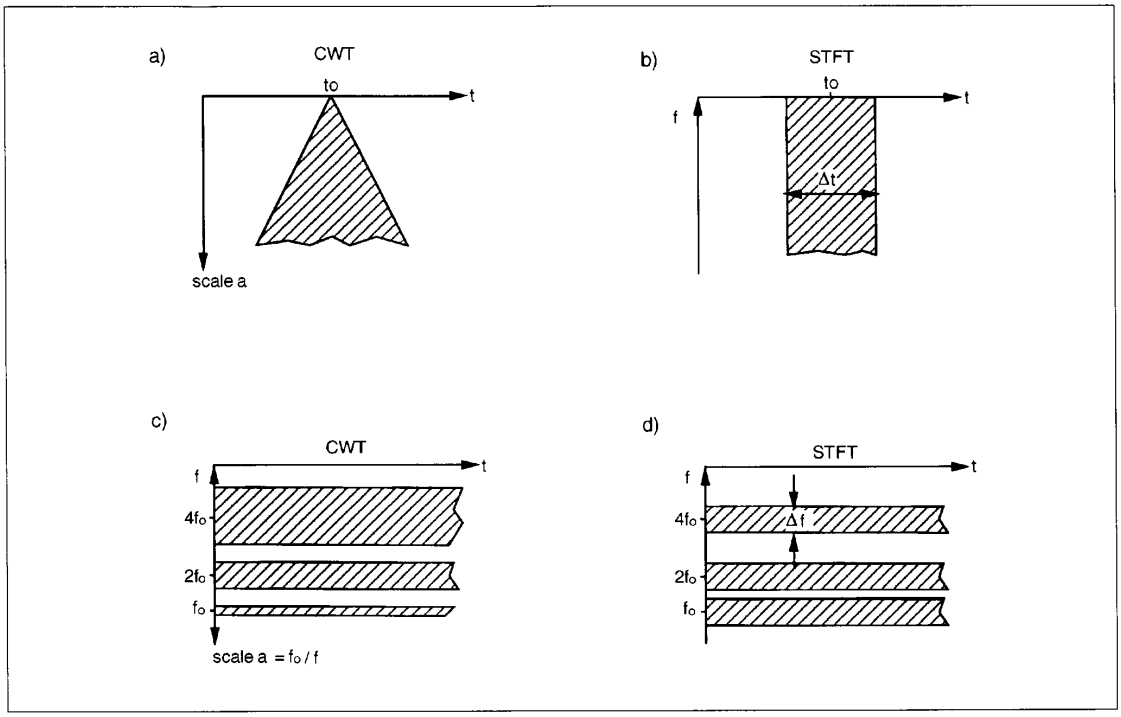


Fig. 4. Regions of influence of a Dirac pulse at $t=t_0$ (a) for the CWT and (b) for the STFT: as well as of three sinusoids (of frequencies f_0, $2f_0$, $4f_0$) for (c) the CWT and (d) the STFT.

plane, associated with measure $\frac{d\tau.da}{a^2}$, and thus expressed in power per frequency unit, like the spectrogram. However, in contrast to the spectrogram, the energy of the signal is here distributed with different resolutions according to Fig. 2b.

Figure 4 illustrates differences between a scalogram and a spectrogram. Figure 4a shows that the influence of the signal's behavior around $t = t_0$ in the analysis is limited to a cone in the time-scale plane; it is therefore very "localized" around t_0 for small scales. In the STFT case, the corresponding region of influence is as large as the extent of the analysis window over all frequencies, as shown in Fig. 4b. Moreover, since the time-scale analysis is logarithmic in frequency, the area of influence of some pure frequency f_0 in the signal increases with f_0 in a scalogram (Fig. 4c), whereas it remains constant in a spectrogram (Fig. 4d).

Both the spectrogram and the scalogram produce a more or less easily interpretable visual two-dimensional representation of signals [GRO89], where each pattern in the time-frequency or time-scale plane contributes to the global energy of the signal. However, such an energy representation has some disadvantages, too. For example, the spectrogram, as well as the scalogram, cannot be inverted in general. Phase information is necessary to reconstruct the signal. Also, since both the spectrogram and the scalogram are bilinear functions of the analyzed signal, cross-terms appear as interferences between patterns in the time-frequency or time-scale plane [KAD91] and this may be undesirable.

In the wavelet case, it has been also shown [GRO89] that the *phase* representation more accurately reveals isolated, local bursts in a signal than the scalogram does (see Box 3).

To illustrate the above points, Fig. 5 shows some examples of spectrograms and scalograms for synthetic signals and a speech signal (see Box 3).

More involved energy representations can be developed for both time-frequency and time-scale [BER88, FLA90, RIO90a], and a link between the spectrogram, the scalogram and the Wigner-Ville distribution can be established (see Box 4).

WAVELET FRAMES AND ORTHONORMAL BASES

Discretization of Time-Scale Parameters

We have seen that the continuously labelled basis functions (wavelets) $h_{a,\tau}(t)$ behave in the wavelet analysis and synthesis just like an orthonormal basis. The following natural question arises: if we appropriately discretize the time-scale parameters a, τ, can we obtain a *true* orthonormal basis? The answer, as we shall see, is that it depends on the choice of the basic wavelet $h(t)$.

There is a natural way to discretize the time-scale parameters a, τ [DAU90a]: since two scales $a_0 < a_1$ roughly correspond to two frequencies $f_0 > f_1$, the wavelet coefficients at scale a_1 can be subsampled at $(f_0/f_1)^{\text{th}}$ the rate of the coefficients at scale a_0, according

BOX 3:
Spectrograms and Scalograms

We present in Fig. 5 spectrograms and scalograms for some synthetic signals and a real signal. The signals are of length 384 samples, and the STFT uses a Gaussian-like window of length $L = 128$ samples. The scalogram is obtained with a Morlet wavelet (a complex sinusoid windowed with a Gaussian envelope) of length from 23 to 363 samples.

The horizontal axis is time in both spectrograms and scalograms. The signal is shown on the top. The vertical axis is frequency in the spectrogram (high frequencies on top) and scale in the scalogram (small scale at the top). Compare these figures with Fig. 4, which indicates the axis system used, and gives the rough behavior for Diracs and sinewaves.

First, Fig. 5.1 shows the analysis of two Diracs and two sinusoids with the STFT and the CWT. Note how the Diracs are well time-localized at high frequencies in the scalogram. Figure 5.2 shows the analysis of three starting sinusoids with different starting times (a low frequency starts first, followed by a medium and a high frequency sinewave). Figure 5.3 shows the transforms of a chirp signal. Again, the transitions are well resolved at high frequencies in the scalogram. Finally, Fig. 5.4 shows the analysis of a segment of speech signal, where the onset of voicing is seen in both representations.

Note that displaying scalograms is sometimes tricky, because parameters like display look-up tables (which map the scalogram value to a grey scale value on the screen) play an important but not always well understood role in the visual impression. Such problems are common in spectrogram displays as well.

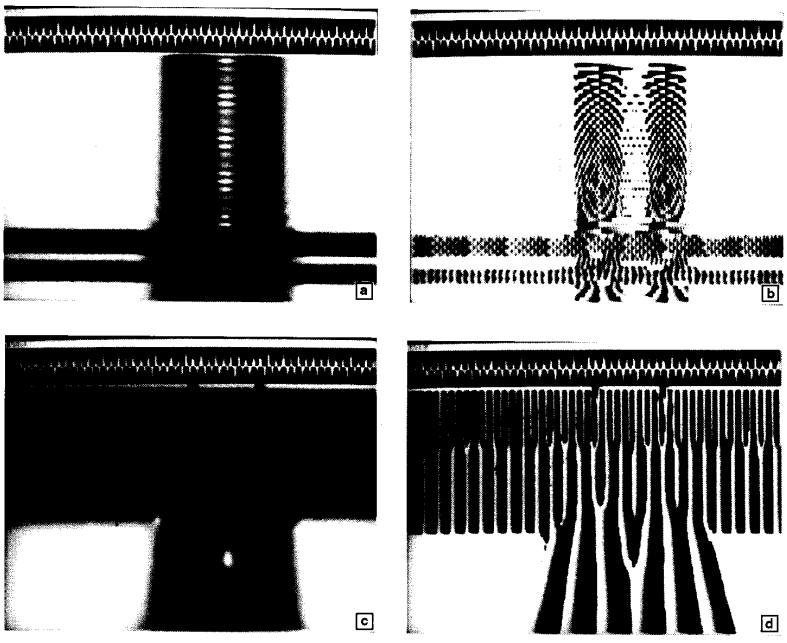


Fig. 5.1. Spectrogram and scalogram for the STFT and CWT analysis of two Dirac pulses and two sinusoids. (a) Magnitude of the STFT. (b) Phase of the STFT. (c) Amplitude of the WT. (d) Phase of the WT.

BOX 3: Spectrograms and Scalograms *(continued)*

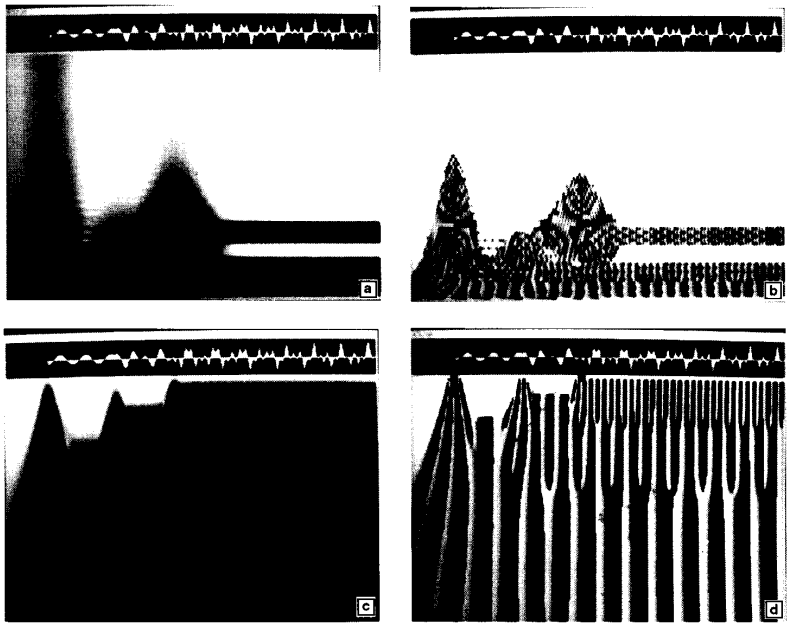


Fig. 5.2. Spectrogram and scalogram for the STFT and CWT analysis of three sinusoids with staggered starting times. The low frequency one comes first, followed by the medium and high frequency ones. (a) Magnitude of the STFT. (b) Phase of the STFT. (c) Amplitude of the WT. (d) Phase of the WT.

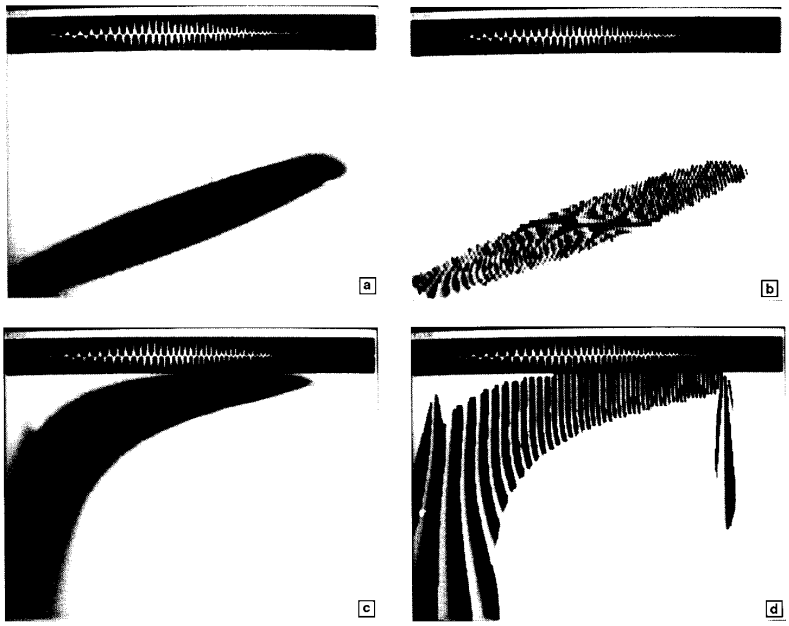


Fig. 5.3. Spectrogram and scalogram for the STFT and CWT analysis of a chirp signal. (a) Magnitude of the STFT. (b) Phase of the STFT. (c) Amplitude of the WT. (d) Phase of the WT.

BOX 3: Spectrograms and Scalograms *(continued)*

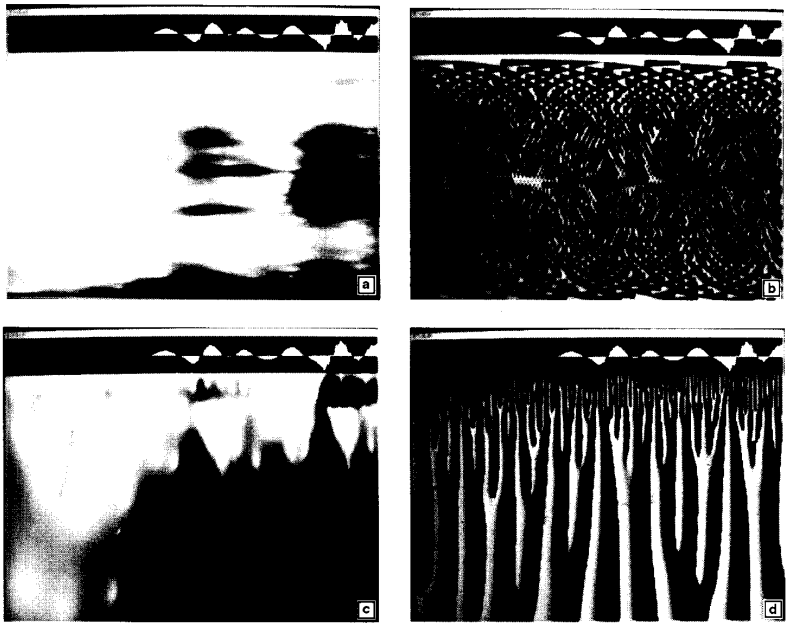


Fig. 5.4. Spectrogram and scalogram for the STFT and CWT analysis of a segment of speech, including onset of voicing. (a) Magnitude of the STFT. (b) Phase of the STFT. (c) Amplitude of the WT. (d) Phase of the WT.

to Nyquist's rule. We therefore choose to discretize the time-scale parameters on the sampling grid drawn in Fig. 7. That is, we have $a = a_0^j$ and $b = k\, a_0^j\, T$, where j and k are integers. The corresponding wavelets are

$$h_{j,k}(t) = a_0^{-j/2}\, h(a_0^{-j}\, t - kT) \tag{9}$$

resulting in wavelet coefficients

$$c_{j,k} = \int x(t)\, h^*_{j,k}(t)\, dt \tag{10}$$

An analogy is the following: assume that the wavelet analysis is like a microscope. First one chooses the magnification, that is, a_0^{-j}. Then one moves to the chosen location. Now, if one looks at very small details, then the chosen magnification is large and corresponds to j negative and large. Then, $a_0^j\, T$ corresponds to small steps, which are used to catch small details. This justifies the choice $b = k\, a_0^j\, T$ in (9).

The reconstruction problem is to find a_0, T, and $h(t)$ such that

$$x(t) \approx c \sum_j \sum_k c_{j,k}\, h_{j,k}(t) \tag{11}$$

where c is a constant that does not depend on the signal (compare with (8)). Evidently, if a0 is close enough to 1 (and if T is small enough), then the wavelet functions are overcomplete. Equation (11) is then still very close to (8) and signal reconstruction takes place within non-restrictive conditions on h(t). On the other hand, if the sampling is sparse, e.g. the computation is done octave by octave ($a_0 = 2$), a true orthonormal basis will be obtained only for very special choices of h(t) [DAU90a, MEY90].

Wavelet Frames

The theory of wavelet frames [DUF52, DAU90a] provides a general framework which covers the two extreme situations just mentioned. It therefore permits one to balance (i) redundancy, i.e. sampling density in Fig. 7, and (ii) restrictions on h(t) for the reconstruction scheme (11) to work. The trade-off is the following: if the redundancy is large (high "oversampling"), then only mild restrictions are put on the basis functions (9). But if the redundancy is small (i.e., close to "critical" sampling), then the basis functions are very constrained.

The idea behind frames [DUF52] is based on the assumption that the linear operator x(t) Ä cj,k, where cj,k is defined by (10), is bounded, with bounded inverse. The family of wavelet functions is then called a frame and is such that the energy of the wavelet coefficients cj,k (sum of their square moduli) relative to that of the signal lies between two positive "frames bounds" A and B,

$$A \cdot E_x \le \sum_{j,k} |\, c_{j,k}\, |^2 \le B \cdot E_x$$

Box 4:
Merging Spectrogram, Scalogram, and Wigner Distribution into a Common Class of Energy Representations

There has been considerable work in extending the spectrogram into more general time-frequency energy distributions TF(τ, f). These all have the basic property of distributing the energy of the signal all over the time-frequency plane, i.e.,

$$\int\int \mathrm{TF}(\tau,f)\, d\tau\, df = \int |x(t)|^2 dt$$

Among them, an alternative to the spectrogram for nonstationary signal analysis is the Wigner-Ville distribution [CLA80, BOU85, FLA89]

$$W_x(\tau,f) = \int x(\tau + \frac{t}{2})\, x^*(\tau - \frac{t}{2})\, e^{-2j\pi ft} dt$$

More generally, the whole class of time-frequency energy distributions has been fully described by Cohen [COH66], [COH89]: they can all be seen as smoothed (or, more precisely, correlated) versions of the Wigner-Ville distribution. The spectrogram is itself recovered when the "smoothing" function is the Wigner-Ville distribution of the analysis window!

A similar situation appears for time-scale energy distributions. For example, the scalogram can be written as [FLA90], [RIO90a]

$$|\mathrm{CWT}(\tau, a)|^2 = \int\int W_x(t,v)\, W_h(\frac{t-\tau}{a}, av)\, dt\, dv$$

i.e., as some 2D "affine" correlation between the signal and the "basic" wavelet's Wigner-Ville distribution. This remarkable formula tells us that there exist strong links between Wavelet Transforms and Wigner-Ville distributions. And, as a matter of fact, it can be generalized to define the most general class of time-scale energy distributions [BER88, FLA90, RIO90a], just as in the time-frequency case.

Figure 6 shows that it is even possible to go continuously from the spectrogram of a given signal to its scalogram [FLA90, RIO90a]. More precisely, starting from the Wigner-Ville distribution, by progressively controlling Gaussian smoothing functions, one goes through a set of energy representations which either tends to the spectrogram if regular two-dimensional smoothing is used, or to the scalogram if "affine" smoothing is used. This property may allow us to decide whether or not we should choose time-scale analysis tools, rather than time-frequency ones for a given problem.

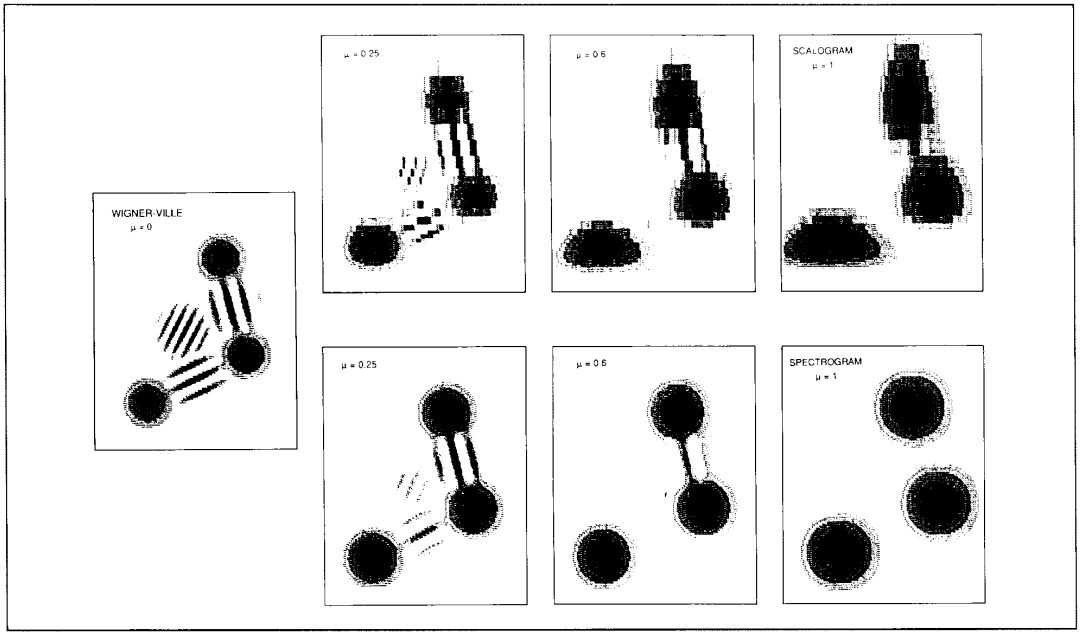


Fig. 6. From spectrograms to scalograms via Wigner-Ville. By controlling the parameter μ (which is a measure of the time-frequency extent of the smoothing function), it is possible to make a full transition between time-scale and time-frequency analyses. Here seven analyses of the same signal (composed of three Gaussian packets) are shown. Note that the best joint time-frequency resolution is attained for the Wigner-Ville distribution, while both spectrogram and scalogram (which can be thought of as smoothed versions of Wigner-Ville) provide reduced cross-term effects compared to Wigner-Ville (after [FLA90, RIO90a]).

where E_x is the energy of the signal $x(t)$.

These frame bounds can be computed from a_0, T and $h(t)$ using Daubechies' formulae [DAU90a]. Interestingly enough, they govern the accuracy of signal reconstruction by (11). More precisely, we have

$$x(t) \approx \frac{2}{A+B} \sum_j \sum_k c_{j,k}\, h_{j,k}(t)$$

with relative SNR greater than $(B/A+1)/(B/A-1)$ (see Fig. 8). The closer A and B, the more accurate the reconstruction. It may even happen that $A=B$ ("tight frame"), in which case the wavelets behave exactly like an orthonormal basis, although they may not even be linearly independent [DAU90a]! The reconstruction formula can also be made exact in the general case if one uses different synthesis functions $h'_{jk}(t)$ (which constitute the *dual frame* of the $h_{jk}(t)$s [DAU90a]).

Introduction to orthogonal wavelet bases

If a tight frame is such that all wavelets $h_{j,k}(t)$ (9) are necessary to reconstruct a general signal, then the wavelets form an *orthonormal* basis of the space of signals with finite energy [HEI90]. Recall that orthonormality means

$$\int h_{j,k}(t)\, h^*_{j',k'}(t)\, dt = \begin{cases} 1 & \text{if } j=j' \text{ and } k=k' \\ 0 & \text{otherwise} \end{cases}$$

An arbitrary signal can then be represented exactly as a weighted sum of basis functions,

$$x(t) = \sum_{j,k} c_{j,k}\, h_{j,k}(t)$$

That is, not only the basis functions $h_{j,k}(t)$ are obtained from a single prototype function $h(t)$ by means of scalings and shifts, but also they form an orthonormal basis. What is most interesting is that there do exist well-behaved functions $h(t)$ that can be used as prototype wavelets, as we shall see below. This is in sharp contrast with the STFT, where, according to the Balian-Low theorem [DAU90a], it is impossible to have orthonormal bases with functions well localized in time and frequency (that is, for which the time-bandwidth product $\Delta t\, \Delta f$ is a finite number).

Recently, the wavelet orthonormal scheme has been extended to synthesis functions $h'_{jk}(t) \neq h_{jk}(t)$, leading to so-called *biorthogonal* wavelet bases [COH90a], [VET90a], [VET90b].

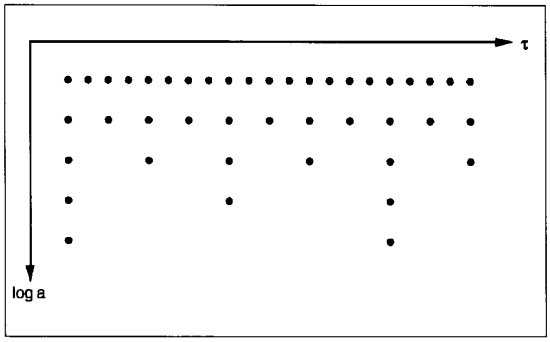


Fig. 7. Dyadic sampling grid in the time-scale plane. Each node corresponds to a wavelet basis function $h_{j,k}(t)$ with scale 2^{-j} and shift $2^{-j}k$.

THE DISCRETE TIME CASE

In this section, we first take a purely discrete-time point of view. Then, through the construction of iterated filter banks, we shall come back to the continuous-time

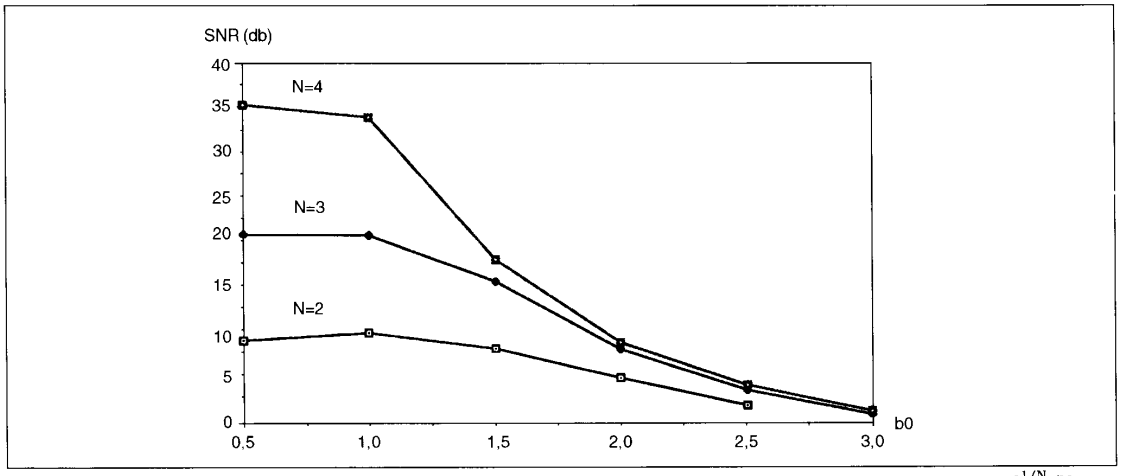


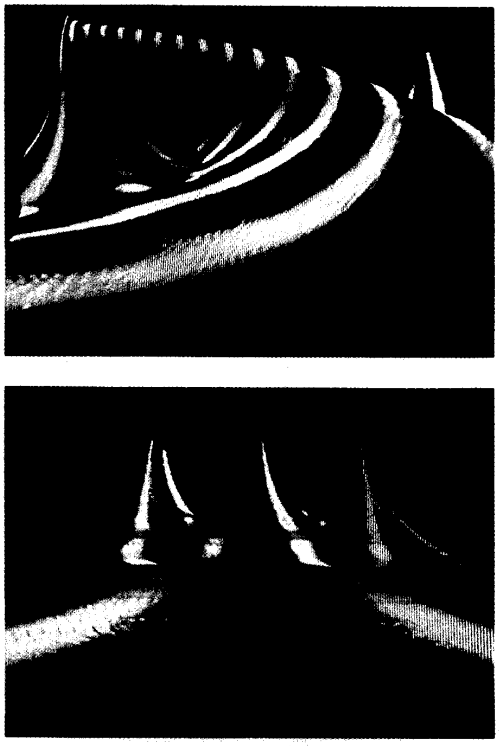
Fig. 8. Reconstruction Signal/Noise Ratio (SNR) error after frame decomposition for different sampling densities $a_0 = 2^{1/N}$ (N = number of voices per octave), $b = a_0^k\, b_0$ (after [DAU90a]). The basic wavelet is the Morlet wavelet (modulated Gaussian) used in [GRO89]. The reconstruction is done "as if" wavelets were orthogonal (see text), and its accuracy grows as N increases and b_0 decreases, i.e. as the density of the sampling grid of Fig. 7 increases. Therefore, redundancy refines the "orthogonal-like" reconstruction.

Three Dimensional Displays of Complex Wavelet Transforms

As seen in Box 3, the wavelet transform using a complex wavelet like the Morlet wavelet (a complex sinusoid windowed by a Gaussian) leads to a complex valued function on the plane.

Phase information is also useful and thus, there is interest in a common display of magnitude and phase. This is possible by using height as magnitude and color as phase, leading to so-called "phasemagrams".

Two examples are shown here: a synthetic chirp in the upper figure (similar to the one in Fig. 5.3); and a triangle function below. In both cases, the discontinuous points are clearly identified at small scales (top of the figure). The chirp has two such points (beginning and end), while the triangle has three. At large scales, these signals look just like a single discontinuity, which is what an observer would indeed see from very far away. For the chirp, the phase cycles with increasing speed, as expected.



Signal analyses with a Morlet wavelet. The display shows magnitude as height and phase as color (phasemagram). The horizontal axis is time. Above) a synthetic chirp signal, with frequency increasing with time. Below) a triangle function.

case and show how to construct orthonormal bases of wavelets for continuous-time signals [DAU88].

In the discrete time case, two methods were developed independently in the late seventies and early eighties which lead naturally to discrete wavelet transforms, namely subband coding [CRI76], [CRO76], [EST77] and pyramidal coding or multiresolution signal analysis [BUR83]. The methods were proposed for coding, and thus, the notion of critical sampling (of requiring a minimum number of samples) was of importance. Pyramid coding actually uses some oversampling, but because it has an easier intuitive explanation, we describe it first.

While the discrete-time case has been thoroughly studied in the filter bank literature in terms of frequency bands (see e.g. [VAI87]), we insist here on notions which are closer to the wavelet point of view, namely those of *scale* and *resolution*. Scale is related to the size of the signal, while resolution is linked to the amount of detail present in the signal (see Box 1 and Fig. 9).

Note that the *scale* parameter in discrete wavelet analysis is to be understood as follows: For large scales, dilated wavelets take "global views" of a *subsampled* signal, while for small scales, contracted wavelets analyze small "details" in the signal.

The Multiresolution Pyramid

Given an original sequence $x(n)$, $n \in \mathbf{Z}$, we derive a lower resolution signal by lowpass filtering with a halfband low-pass filter having impulse response $g(n)$. Following Nyquist's rule, we can subsample by two (drop every other sample), thus doubling the scale in the analysis. This results in a signal $y(n)$ given by

$$y(n) = \sum_{k=-\infty}^{+\infty} g(k)\, x(2n - k)$$

The resolution change is obtained by the lowpass filter (loss of high frequency detail). The scale change is due to the subsampling by two, since a shift by two in the original signal $x(n)$ results in a shift by one in $y(n)$.

Now, based on this lowpass and subsampled version of $x(n)$, we try to find an approximation, $a(n)$, to the original. This is done by first upsampling $y(n)$ by two (that is, inserting a zero between every sample) since we need a signal at the original scale for comparison.

$$y'(2n) = y(n), \quad y'(2n+1) = 0$$

Then, $y'(n)$ is interpolated with a filter with impulse response $g'(n)$ to obtain the approximation $a(n)$,

$$a(n) = \sum_{k=-\infty}^{\infty} g'(k)\, y'(n - k)$$

Note that if $g(n)$ and $g'(n)$ were perfect halfband filters (having a frequency passband equal to 1 over the normalized frequency range $-\pi/2$, $\pi/2$ and equal to 0 elsewhere), then the Fourier transform of a(n) would be

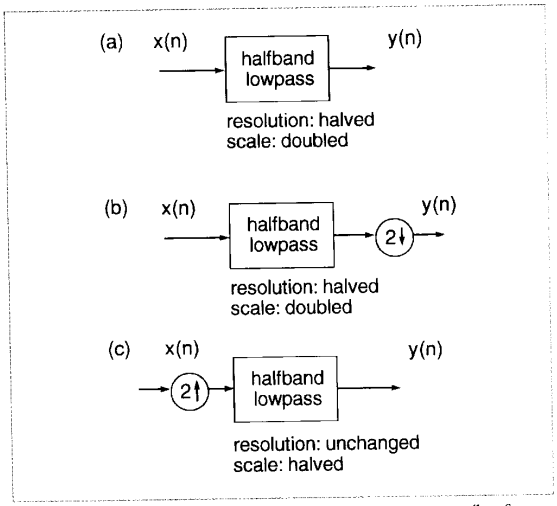


Fig. 9. Resolution and scale changes in discrete time (by factors of 2). Note that the scale of signals is defined as in geographical maps. (a) Halfband lowpass filtering reduces the resolution by 2 (scale is unchanged). (b) Halfband lowpass filtering followed by subsampling by 2 doubles the scale (and halves the resolution as in (a)). (c) Upsampling by 2 followed by halfband lowpass filtering halves the scale (resolution is unchanged).

equal to the Fourier transform of $x(n)$ over the frequency range $(-\pi/2, \pi/2)$ while being equal to zero elsewhere. That is, $a(n)$ would be a perfect halfband lowpass approximation to $x(n)$.

Of course, in general, $a(n)$ is not going to be equal to $x(n)$ (in the previous example, $x(n)$ would have to be a halfband signal). Therefore, we compute the difference between $a(n)$ (our approximation based on $y(n)$) and $x(n)$,

$$d(n) = x(n) - a(n)$$

It is obvious that $x(n)$ can be reconstructed by adding $d(n)$ and $a(n)$, and the whole process is shown in Fig. 10. However, there has to be some redundancy, since a signal with sampling rate f_s is mapped into two signals $d(n)$ and $y(n)$ with sampling rates f_s and $f_s/2$, respectively.

In the case of a perfect halfband lowpass filter, it is clear that $d(n)$ contains exactly the frequencies above $\pi/2$ of $x(n)$, and thus, $d(n)$ can be subsampled by two as well without loss of information. This hints at the fact that critically sampled schemes must exist.

The separation of the original signal $x(n)$ into a coarse approximation $a(n)$ plus some additional detail contained in $d(n)$ is conceptually important. Because of the resolution change involved (lowpass filtering followed by subsampling by two produces a signal with half the resolution and at twice the scale of the original), the above method and related ones are part of what is called Multiresolution Signal Analysis [ROS84] in computer vision.

The scheme can be iterated on $y(n)$, creating a hierarchy of lower resolution signals at lower scales. Because of that hierarchy and the fact that signals become shorter and shorter (or images become smaller and smaller), such schemes are called signal or image pyramids [BUR83].

Subband Coding Schemes

We have seen that the above system creates a redundant set of samples. More precisely, one stage of a pyramid decomposition leads to both a half rate low resolution signal and a full rate difference signal, resulting in an increase in the number of samples by 50%. This oversampling can be avoided if the filters $g(n)$ and $g'(n)$ meet certain conditions [VET90b].

We now look at a different scheme instead, where no such redundancy appears. It is the so-called subband coding scheme first popularized in speech compression [CRI76, CRO76, EST77]. The lowpass, subsampled approximation is obtained exactly as explained above, but, instead of a difference signal, we compute the "added detail" as a highpass filtered version of $x(n)$ (using a filter with impulse response $h(n)$), followed by subsampling by two. Intuitively, it is clear that the "added detail" to the lowpass approximation has to be a highpass signal, and it is obvious that if $g(n)$ is an ideal halfband lowpass filter, then an ideal halfband highpass filter $h(n)$ will lead to a perfect representation of the original signal into two subsampled versions.

This is exactly one step of a wavelet decomposition using $sin(x)/x$ filters, since the original signal is mapped into a lowpass approximation (at twice the scale) and

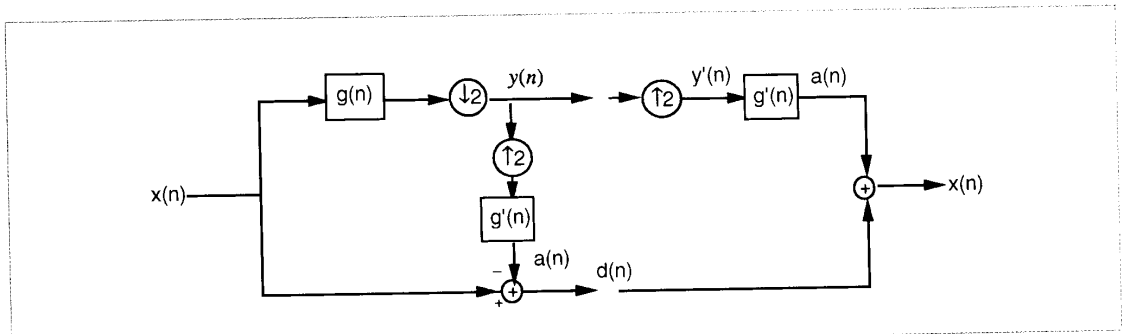


Fig. 10. Pyramid scheme. Derivation of a lowpass, subsampled approximation y(n), from which an approximation a(n) to x(n) is derived by upsampling and interpolation. Then, the difference between the approximation a(n) and the original x(n) is computed as d(n). Perfect reconstruction is simply obtained by adding a(n) back.

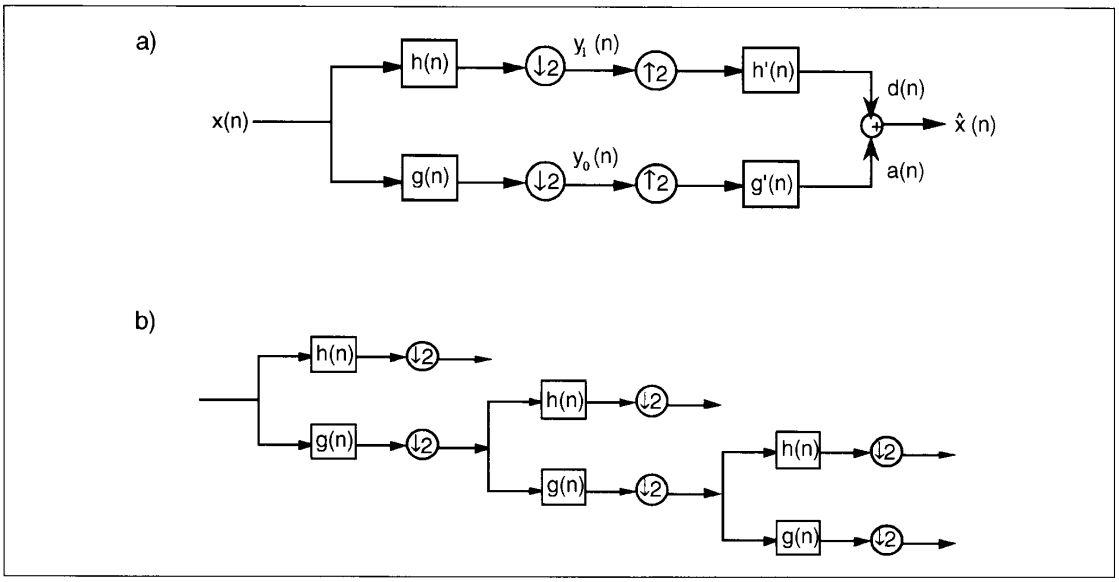


Fig. 11. Subband Coding scheme. (a) Two subsampled approximations, one corresponding to low and the other to high frequencies, are computed. The reconstructed signal is obtained by re-interpolating the approximations and summing them. The filters on the left form an analysis filter bank, while on the right is a synthesis filter bank. (b) Block diagram (Filter Bank tree) of the Discrete Wavelet Transform implemented with discrete-time filters and subsampling by two. The frequency resolution is given in Fig. 3b.

an added detail signal (also at twice the scale). In particular, using these ideal filters, the discrete version is identical to the continuous wavelet transform.

What is more interesting is that it is not necessary to use ideal (that is, impractical) filters, and yet $x(n)$ can be recovered from its two filtered and subsampled versions which we now call $y_0(n)$ and $y_1(n)$. To do so, both are upsampled and filtered by $g'(n)$ and $h'(n)$ respectively, and finally added together, as shown in Fig. 11a. Now, unlike the pyramid case, the reconstructed signal (which we now call $\hat{x}(n)$) is not identical to $x(n)$, unless the filters meet some specific constraints. Filters that meet these constraints are said to have *perfect reconstruction* property, and there are a number of papers investigating the design of perfect reconstruction filter banks [MIN85, SMI86, VAI88, VET86].

The easiest case to analyze appears when the analysis and synthesis filters in Fig. 11a are identical (within time-reversal) and when perfect reconstruction is achieved (that is, $\hat{x}(n) = x(n)$, within a possible shift). Then it can be shown that the subband analysis/synthesis corresponds to a decomposition onto an orthonormal basis, followed by a reconstruction which amounts to summing up the orthogonal projections. We will assume FIR filters in the following. Then, it turns out that the highpass and lowpass filters are related by

$$h(L-1-n) = (-1)^n g(n) \tag{12}$$

where L is the filter length (which has to be even). Note that the modulation by $(-1)^n$ transforms indeed the lowpass filter into a highpass one.

Now, the filter bank in Fig. 11a, which computes convolutions followed by subsampling by two, evaluates inner products of the sequence $x(n)$ and the sequences $\{g(-n+2k), h(-n+2l)\}$ (the time reversal comes from the convolution, which reverses one of the sequences). Thus

$$y_0(k) = \sum_n x(n)\, g(-n+2k)$$
$$y_1(k) = \sum_n x(n)\, h(-n+2k)$$

Because the filter impulse responses form an orthonormal set, it is very simple to reconstruct $x(n)$ as

$$x(n) = \sum_{k=-\infty}^{\infty} \left[y_0(k)\, g(-n+2k) + y_1(k)\, h(-n+2k) \right] \tag{13}$$

that is, as a weighted sum of the orthogonal impulse responses, where the weights are the inner products of the signal with the impulse responses. This is of course the standard expansion of a signal into an orthonormal basis, where the resynthesis is the sum of the orthogonal projections (see *Introduction to orthogonal wavelet bases* above).

From (12), (13) it is also clear that the synthesis filters are identical to the analysis filters within time reversal.

Such orthogonal perfect reconstruction filter banks have been studied in the digital signal processing literature, and the orthonormal decomposition we just indicated is usually referred to as a "paraunitary" or "lossless" filter bank [VAI89]. An interesting property of such filter banks is that they can be written in lattice form [VAI88], and that the structure and properties can be extended to more than two channels [VAI87, VAI89,

VET89]. More general perfect reconstruction (biorthogonal) filter banks have also been studied (see e.g. [VET86, VET90b, COH90a]). It has been also noticed [MAL89b, SHE90, RIO90b] that filter banks arise naturally when implementing the CWT.

Note that we have assumed linear processing throughout. If non-linear processing is involved (like quantization), the oversampled nature of the pyramid scheme described in the preceding section may actually lead to greater robustness.

The Discrete Wavelet Transform

We have shown how to decompose a sequence $x(n)$ into two subsequences at half rate, or half resolution, and this by means of "orthogonal" filters (orthogonal with respect to even shifts). Obviously, this process can be iterated on either or both subsequences. In particular, to achieve finer frequency resolution at lower frequencies (as obtained in the continuous wavelet transform), we iterate the scheme on the lower band only. If $g(n)$ is a good halfband lowpass filter, $h(n)$ is a good halfband highpass filter by (12). Then, one iteration of the scheme on the first lowband creates a new lowband that corresponds to the lower quarter of the frequency spectrum. Each further iteration halves the width of the lowband (increases its frequency resolution by two), but due to the subsampling by two, its time resolution is halved as well. At each iteration, the current high band portion corresponds to the difference between the previous lowband portion and the current one, that is, a passband. Schematically, this is equivalent to Fig. 11b, and the frequency resolution is as in Fig. 3b.

An important feature of this discrete algorithm is its relatively low complexity. Actually, the following somewhat surprising result holds: independent of the depth of the tree in Fig. 11b, the complexity is linear in the number of input samples, with a constant factor that depends on the length of the filter. The proof is straightforward. Assume the computation of the first filter bank requires C_0 operations per input sample (C_0 is typically of the order of L). Then, the second stage requires also C_0 operations per sample of its input, but, because of the subsampling by two, this amounts to $C_0/2$ operations per sample of the input signal. Therefore, the total complexity is bounded by

$$C_{total} = C_0 + \frac{C_0}{2} + \frac{C_0}{4} + \ldots < 2C_0$$

which demonstrates the efficiency of the discrete wavelet transform algorithm and shows that it is independent of the number of octaves that one computes. This bounded complexity had been noticed in the multirate filtering context [RAM88]. Further developments can be found in [RIO91a]. Note that a possible drawback is that the delay associated with such an iterated filter bank grows exponentially with the number of stages.

Iterated Filters and Regularity

There is a major difference between the discrete scheme we have just seen and the continuous time wavelet transform. In the discrete time case, the role of the wavelet is played by the highpass filter $h(n)$ and the cascade of subsampled lowpass filters followed by a highpass filter (which amounts to a bandpass filter). These filters, which correspond roughly to octave band filters, unlike in the continuous wavelet transform, are not exact scaled versions of each other. In particular, since we are in discrete time, scaling is not as easily defined, since it involves interpolation as well as time expansion.

Nonetheless, under certain conditions, the discrete system converges (after a certain number of iterations) to a system where subsequent filters are scaled versions of each other. Actually, this convergence is the basis for the construction of continuous time compactly supported wavelet bases [DAU88].

Now, we would like to find the equivalent filter that corresponds to the lower branch in Fig. 11b, that is the iterated lowpass filter. It will be convenient to use z-transforms of filters, e.g. $G(z) = \sum_n g(n)\, z^{-n}$ in the following. It can be easily checked that subsampling by two followed by filtering with $G(z)$ is equivalent to filtering with $G(z^2)$ followed by the subsampling (z^2 inserts zeros between samples of the impulse response, which are removed by the subsequent subsampling). That is, the first two steps of lowpass filtering can be replaced by a filter with z-transform $G(z)\cdot G(z^2)$, followed by subsampling by 4. More generally, calling $G^i(z)$ the equivalent filter to i stages of lowpass filtering and subsampling by two (that is, a total subsampling by 2^i), we obtain

$$G^i(z) = \prod_{l=0}^{i-1} G(z^{2^l}) \qquad (14)$$

Call its impulse response $g^i(n)$.

As i infinitely increases, this filter becomes infinitely long. Instead, consider a function $f^i(x)$ which is piecewise constant on intervals of length $1/2^i$ and has value $2^{i/2}\, g^i(n)$ in the interval $[n/2^i, (n+1)/2^i]$. That is, $f^i(x)$ is a staircase function with the value given by the samples of $g^i(n)$ and intervals which decrease as 2^{-i}. It can be verified that the function is supported on the interval $[0, L-1]$, where L is the length of the filter $g(n)$. Now, for i going to infinity, $f^i(x)$ can converge to a continuous function $g_c(x)$, or a function with finitely many discontinuities, even a fractal function, or not converge at all (see Box 5).

A necessary condition for the iterated functions to converge to a continuous limit is that the filter $G(z)$ should have a sufficient number of zeros at $z = -1$, or half sampling frequency, so as to attenuate repeat spectra [DAU88, DAU90b, RIO91b]. Using this condition, one can construct filters which are both orthogonal and converge to continuous functions with compact support. Such filters are called *regular*, and examples can be found in [DAU88, COH90a, DAU90b, RIO90b, VET90b]. Note that the above condition can be interpreted as a *flatness* condition on the spectrum of $G(z)$ at half sampling frequency. In fact, it can be shown

Box 5: Regular Scaling Filters

It is well known that the structure of computations in a Discrete Wavelet Transform and in an octave-band filter bank are identical. Therefore, besides the different views and interpretations that have been given to them, the main difference lies in the filter design. Wavelet filters are chosen so as to be *regular*. Recall that this means (with the same notation as used in the main text sections on iterated filters), that the piecewise constant function associated with the discrete wavelet sequence $h_j(n)$ of z-transform $G^j(z)H(z^{2^j})$ converges (e.g. pointwise), as j indefinitely increases, to a regular limit function $h_c(x)$. Equivalently, the piecewise constant function associated with the discrete "scaling" sequence $g_j(n)$ of z-transform $G^j(z)$ converges to a regular limit function $g_c(x)$. By "regular" we mean that the continuous-time wavelet $h_c(x)$ (or the scaling function $g_c(x)$) is at least continuous, or better, once or twice continuously differentiable. The regularity order is the number of times $h_c(x)$ (or $g_c(x)$) is continuously differentiable. Figures 12a and 12b show two examples, one where $g_c(x)$ is almost three times continuously differentiable and another where $g_j(n)$ diverges with fractal behavior.

Note that there are a number of classical filters, designed for two-band filter banks, which, unlike wavelet filters, are *not* regular. Figures 12c through 12f show two well-known examples: a Johnston filter [JOH80], and a Smith and Barnwell filter [SMI86]. The latter allows perfect reconstruction, while the former does not. Figure 12d shows that the Smith and Barnwell discrete sequences $h_j(n)$ do not tend to regular limit functions, but rather diverge. This is not surprising since the necessary condition that the low-pass filter has a zero at half the sampling frequency is violated (although this filter has 40 dB attenuation in the stop band [SMI86]). This eventually results, when j increases, in small, but rapid oscillations in $h_j(n)$. As for the Johnston filter (Figs. 12e and 12f), it can be shown that the wavelet limit function is continuous but not differentiable.

For wavelet filters, the more regular the limit function, the faster the convergence to this limit [RIO91b] — and in practice the convergence is very fast. This justifies the study of the limit $h_c(x)$, which is almost attained after a few octaves of a logarithmic decomposition. Since an error in a wavelet coefficient (due e.g. to quantization) results, after reconstruction, in an overall error proportional to a discrete wavelet $h_j(n)$, regularity seems a nice property, e.g., to avoid visible distortion on a reconstructed image [ANI90].

From equations (12), (15) and (16), the knowledge of $g(n)$ suffices to determine the limit $h_c(x)$. Several methods have been developed to estimate the regularity order of $h_c(x)$ from the coefficients $g(n)$. Most are based on Fourier transform techniques [DAU88, COH90b]. Recently, time-domain techniques have been developed which provide optimal estimates [DAU90c, RIO91b].

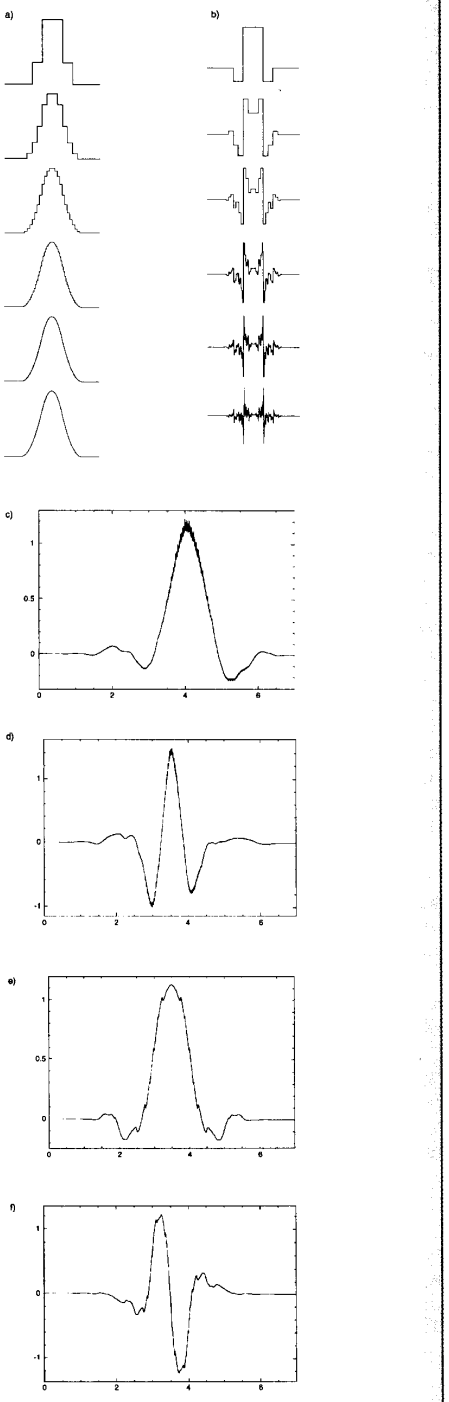


Fig. 12. (a) Iterated low-pass filter $g_j(n)$ with $g(n) = (1,3,3,1)$, converges to a regular, smooth function. (b) Iterated low-pass filter with $g(n) = (-1,3,3,-1)$ converges to a fractal function (see text). (c) Smith and Barnwell 8-tap lowpass filter [SMI86], iterated 8 times. Divergence occurs, due to rapid oscillations in the temporal waveform. (d) Corresponding continuous-time wavelet (after [RIO90b]). (e) Johnston 8-tap lowpass filter [JOH80], iterated 8 times. The limit function is not differentiable. (f) Corresponding continuous-time wavelet.

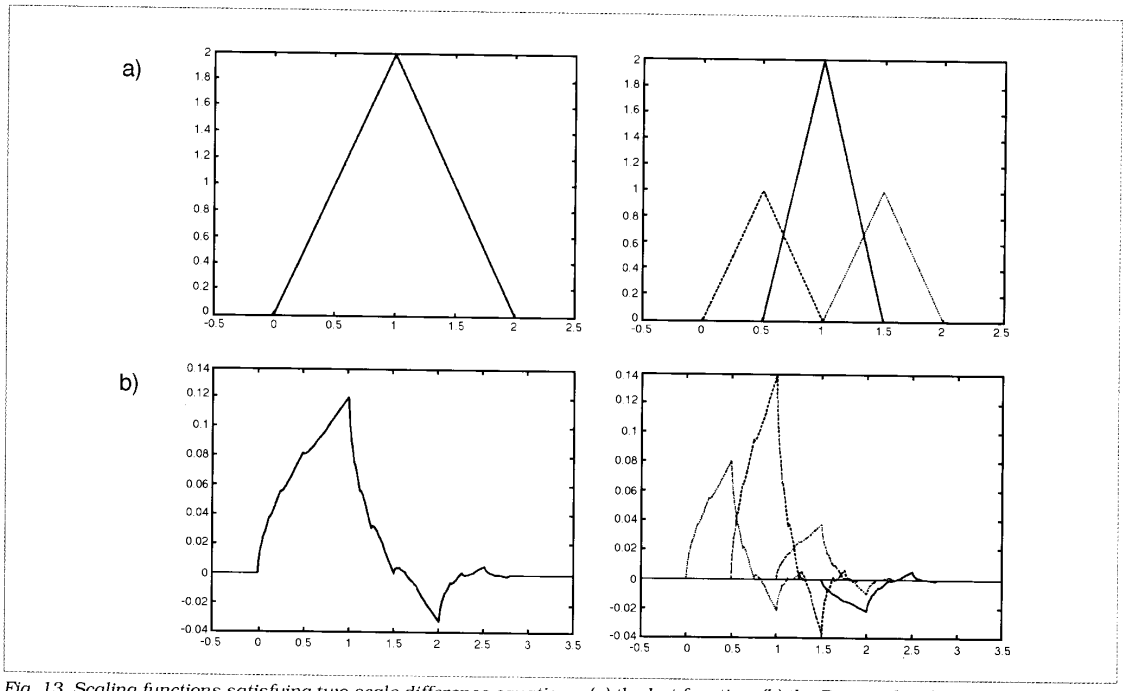


Fig. 13. Scaling functions satisfying two-scale difference equations. (a) the hat function. (b) the D_4 wavelet obtained from a 4-tap regular filter by Daubechies.

[AKA90], [SHE90] that the well-known Daubechies orthonormal filters [DAU88] are deduced from "maximally flat" low-pass filters [HER71]. Note that there are many other choices that behave very differently in terms of phase, selectivity in frequency, and other criteria (see e.g. [DAU90b]). An important issue related to regular filter design is the derivation of simple estimates for the regularity order (see Box 5).

It is still not clear whether regular filters are most adapted to coding schemes [ANI90]. The minimal regularity order necessary for good coding performance of discrete wavelet transform schemes, if needed at all, is also not known and remains a topic for future investigation.

Scaling Functions and Wavelets Obtained from Iterated Filters

Recall that $g_c(x)$ is the final function to which $f^i(x)$ converges. Because it is the product of lowpass filters, the final function is itself lowpass and is called a "*scaling function*" because it is used to go from a fine scale to a coarser scale. Because of the product (14) from which the scaling function is derived, $g_c(x)$ satisfies the following two scale difference equation [DAU90c]:

$$g_c(x) = \sum_{n=-\infty}^{\infty} g(n)\, g_c(2x-n) \qquad (15)$$

Figure 13 shows two such examples. The second one is based on the 4-tap Daubechies filter which is regular and orthogonal to its even translates [DAU88].

So far, we have only discussed the iterated lowpass and its associated scaling function. However, from Fig. 11b, it is clear that a bandpass filter is obtained in the same way, except for a final highpass filter. Therefore, in a fashion similar to (15), the wavelet $h_c(x)$ is obtained as

$$h_c(x) = \sum_{n=-\infty}^{\infty} h(n)\, g_c(2x-n) \qquad (16)$$

that is, it also satisfies a two scale equation.

Now, if the filters $h(n)$ and $g(n)$ form an orthonormal set with respect to even shifts, then the functions $g_c(x-l)$ and $h_c(x-k)$ form an orthonormal set (see Box 6). Because they also satisfy two scale difference equations, it can be shown [DAU88] that the set $h_c(2^{-i}x-k)$, $i,k \in \mathbf{Z}$, forms an orthonormal basis for the set of square integrable functions.

Figure 14 shows two scales and shifts of the 4-tap Daubechies wavelet [DAU88]. While it might not be obvious from the figure, these functions are orthogonal to each other, and together with all scaled and translated versions, they form an orthonormal basis.

Figure 15 shows an orthogonal wavelet based on a length-18 regular filter. It is obviously a much smoother function (actually, it possesses 3 continuous derivatives).

Finally, Fig. 16 shows a biorthogonal set of linear phase wavelets, where the analysis wavelets are orthogonal to the synthesis wavelets. These were obtained from a biorthogonal linear phase filter bank with length-18 regular filters [VET90a, VET90b].

We have shown how regular filters can be used to

Box 6: Multiresolution analysis

The concept of multiresolution approximation of functions was introduced by Meyer and Mallat [MAL89a, MAL89c, MEY90] and provides a powerful framework to understand wavelet decompositions. The basic idea is that of successive approximation, together with that of "added detail" as one goes from one approximation to the next, finer one. We here give the intuition behind the construction.

Assume we have a ladder of spaces such that:

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

with the property that if $f(x) \in V_i$ then $f(x-2^{-i}k) \in V_i$, $k \in \mathbf{Z}$, and $f(2x) \in V_{i-1}$. Call W_i the orthogonal complement of V_i in V_{i-1}. This is written

$$V_{i-1} = V_i \stackrel{\perp}{\oplus} W_i \qquad \text{(B6.1)}$$

Thus, W_i contains the "detail" necessary to go from V_i to V_{i-1}. Iterating (B6.1), one has

$$V_{i-1} = W_i \stackrel{\perp}{\oplus} W_{i+1} \stackrel{\perp}{\oplus} W_{i+2} \stackrel{\perp}{\oplus} W_{i+3} \stackrel{\perp}{\oplus} \ldots \qquad \text{(B6.2)}$$

that is, a given resolution can be attained by a sum of added details.

Now, assume we have an orthonormal basis for V_0 made up of a function $g_c(x)$ and its integer translates. Because $V_0 \in V_{-1}$, $g_c(x)$ can be written in terms of the basis in V_{-1}, i.e., (15) is satisfied:

$$g_c(x) = \sum_n c_n \, g_c(2x-n)$$

Then it can be verified that the function $h_c(x)$ (16) (with the relation (12)) and its integer translates form an orthonormal basis for W_0. And, because of (B6.2), $h_c(x)$ and its scaled and translated versions form a wavelet basis [MAL89a, MAL89c, MEY90].

The multiresolution idea is now very intuitive. Assume we have an approximation of a signal at a resolution corresponding to V_0. Then a better approximation is obtained by adding the details corresponding to W_0, that is, the projection of the signal in W_0. This amounts to a weighted sum of wavelets at that scale. Thus, by iterating this idea, a square integrable signal can be seen as the successive approximation or weighted sum of wavelets at finer and finer scale.

generate wavelet bases. The converse is also true. That is, orthonormal sets of scaling functions and wavelets can be used to generate perfect reconstruction filter banks [DAU88, MAL89a, MAL89c].

Extension of the wavelet concept to multiple dimensions, which is useful, e.g. for image coding, is shown in Box 7.

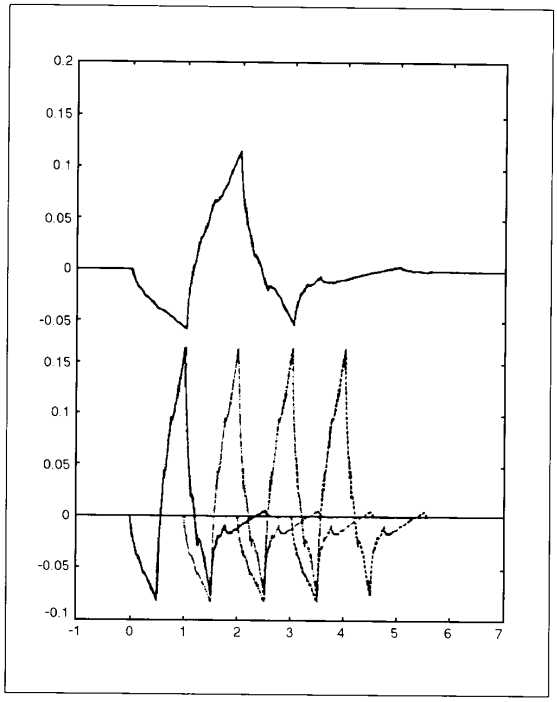


Fig. 14. Two scales of the D_4 wavelet and shifts. This set of functions is orthogonal.

APPLICATIONS OF WAVELETS IN SIGNAL PROCESSING

From the derivation of the wavelet transform as an alternative to the STFT, it is clear that one of the main applications will be in non-stationary signal analysis. While conceptually, the CWT is a classical constant-Q analysis, its simple definition (based on a single function rather than multiple filters) allows powerful analytical derivations and has already lead both to new insights and new theoretical results [WAV89].

Applications of wavelet decompositions in numerical analysis, e.g. for solving partial differential equations, seem very promising because of the "zooming" property which allows a very good representation of discontinuities, unlike the Fourier transform [BEY89].

Perhaps the biggest potential of wavelets has been claimed for signal compression. Since discrete wavelet transforms are essentially subband coding systems, and since subband coders have been successful in speech and image compression, it is clear that wavelets will find immediate application in compression problems. The only difference with traditional subband coders is the fact that filters are designed to be regular (that is, they have many zeroes at $z = 0$ or $z = \pi$). Note that although classical subband filters are not regular (see Box 5 and Fig. 12), they have been designed to have good stopbands and thus are close to being "regular", at least for the first few octaves of subband decomposition.

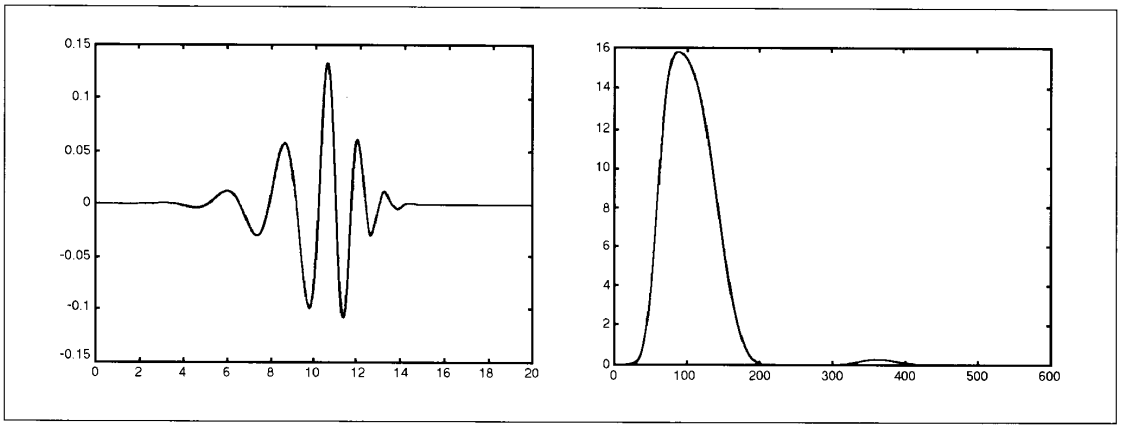


Fig. 15. Orthonormal wavelet generated from a length-18 regular filter [DAU88]. The time function is shown on the left and the spectrum is on the right.

It is therefore clear that drastic improvements of compression will not be achieved so easily simply because wavelets are used. However, wavelets bring new ideas and insights. In this respect, the use of wavelet decompositions in connection with *other* techniques (like vector quantization [ANI90] or multiscale edges [MAL89d]) are promising compression techniques which make use of the elegant theory of wavelets.

New developments, based on wavelet concepts, have already appeared. For example, statistical signal processing using wavelets is a promising field. Multiscale models of stochastic processes [BAS89], [CHO91], and analysis and synthesis of $1/f$ noise [GAC91], [WOR90] are examples where wavelet analysis has been successful. "Wavelet packets" [WIC89], which correspond to arbitrary adaptive tree-structured filter banks, are another promising example.

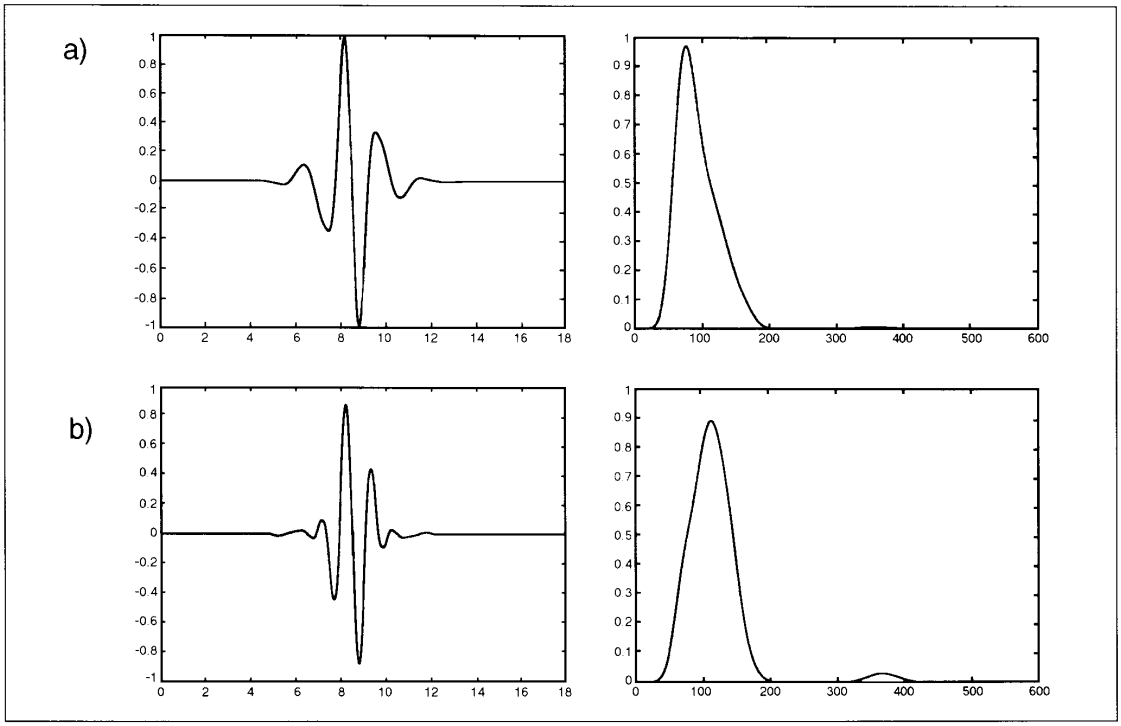


Fig. 16. Biorthogonal wavelets generated from 18-tap regular filters [VET90b]. (a) Analysis wavelet. (b) Synthesis wavelet. The time function is shown on the left and the spectrum is on the right.

Box 7: Multidimensional filter banks and wavelets

In order to apply wavelet decompositions to multidimensional signals (e.g., images), multidimensional extensions of wavelets are required. An obvious way to do this is to use "separable wavelets" obtained from products of one-dimensional wavelets and scaling functions [MAL89a, MAL89c, MEY90]. Let us consider the two-dimensional case for its simplicity. Take a scaling function $g_c(x)$ (15) and a wavelet $h_c(x)$ (16). One can construct for two-dimensional functions :

$$g_c(x,y) = g_c(x) \cdot g_c(y)$$

$$h_c^{(1)}(x,y) = g_c(x) \cdot h_c(y)$$

$$h_c^{(2)}(x,y) = h_c(x) \cdot g_c(y)$$

$$h_c^{(3)}(x,y) = h_c(x) \cdot h_c(y)$$

which are orthogonal to each other with respect to integer shifts (this follows from the orthogonality of the one dimensional component). The function $g_c(x,y)$ is a separable two-dimensional scaling function (that is, a lowpass filter) while the functions $h_c^{(i)}(x,y)$ are "wavelets". The set $\{h_c^{(i)}(2^j x-k,\ 2^j x-l),\ i = 1,2,3$ and $j,k,l \in \mathbf{Z}\}$ forms an orthonormal basis for square integrable functions over $\mathbf{R}^2$. This solution corresponds to a separable two-dimensional filter bank with subsampling by 2 in each dimension, that is, overall subsampling by 4 (see Fig. 17).

More interesting (that is, non-trivial) multidimensional wavelet schemes are obtained when non-separable subsampling is used [KOV92]. For example, a non-separable subsampling by 2 of a double indexed signal $x(n_1, n_2)$ is obtained by retaining only samples satisfying:

$$\begin{pmatrix} n_1 \\ n_2 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}, \quad u_1, u_2 \in \mathbf{Z} \qquad \text{(B7.1)}$$

The resulting points are located on a so-called quincunx sublattice of $\mathbf{Z}^2$. Now, one can construct a perfect reconstruction filter bank involving such subsampling because it resembles its one-dimensional counterpart [KOV92]. The subsampling rate is 2 (equal to the determinant of the matrix in (B7.1)), and the filter bank has 2 channels. Iteration of the filter bank on the lowpass branch (see Fig. 18) leads to a discrete wavelet transform, and if the filter is regular (which now depends on the matrix representing the lattice [KOV92]), one can construct non-separable wavelet bases for square integrable functions over $\mathbf{R}^2$ with a resolution change by 2 (and not 4 as in the separable case). An example scaling function is pictured in Fig. 19.

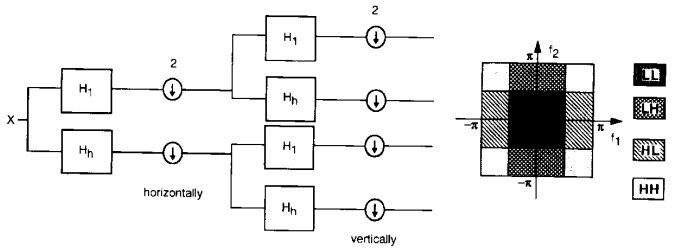


Fig. 17. Separable two-dimensional filter bank corresponding to a separable wavelet basis with resolution change by 4 (2 in each dimension). The partition of the frequency plane is indicated on the right. H_l and H_h stand for low-pass and high-pass filter, respectively.

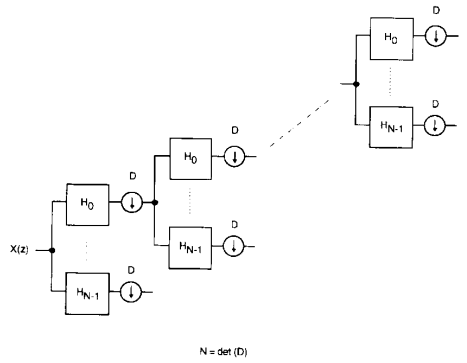


Fig. 18. Iteration of a non-separable filter bank based on non-separable subsampling. This construction leads to non-separable wavelets.

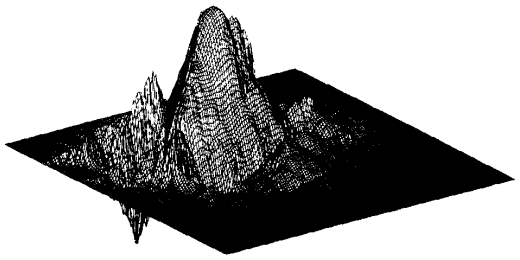


Fig. 19. Two-dimensional non-separable orthonormal scaling function [KOV92] (orthogonality is with respect to integer shifts). The resolution change is by 2 ($\sqrt{2}$ in each dimension). The matrix used for the subsampling is the one given in (B7.1).

CONCLUSION

We have seen that the Short-Time Fourier Transform and the Wavelet Transform represent alternative ways to divide the time-frequency (or time-scale) plane. Two major advantages of the Wavelet Transform are that it can zoom in to time discontinuities and that orthonormal bases, localized in time and frequency, can be constructed. In the discrete case, the Wavelet Transform is equivalent to a logarithmic filter bank, with the added constraint of regularity on the lowpass filter.

The theory of wavelets can be seen as a common framework for techniques that had been developed independently in various fields. This conceptual unification furthers the understanding of the mechanisms involved, quantifies trade-offs, and points to new potential applications. A number of questions remain open, however, and will require further investigations (e.g., what is the "optimal" wavelet for a particular application?).

While some see wavelets as a very promising brand new theory [CIP90], others express some doubt that it represents a major breakthrough. One reason for skepticism is that the concepts have been around for some time, under different names. For example, wavelet transforms can be seen as constant-Q analysis [YOU78], wide-band cross-ambiguity functions [SPE67, AUS90], Frazier-Jawerth transforms [FRA86], perfect reconstruction octave-band filter banks [MIN85, SMI86], or a variation of Laplacian pyramid decomposition [BUR83], [BUR89]!

We think that the interest and merit of wavelet theory is to unify all this into a common framework, thereby allowing new ideas and developments.

ACKNOWLEDGMENTS

The authors would like to thank C. Herley for many useful suggestions and for generating the continuous STFT and WT plots; and Profs. F. Boudreaux-Bartels, M.J.T. Smith and Dr. P. Duhamel for useful suggestions on the manuscript. We thank B. Shakib (IBM) for creating the three-dimensional color rendering of phase and magnitude of wavelet transforms (so-called "phasemagrams") used in the cover picture and elsewhere; C.A. Pickover (IBM) is thanked for his 3D display software and J.L. Mannion for his help on software tools. The second author would like to acknowledge support by NSF under grants ECD-88-11111 and MIP-90-14189.

Olivier Rioul was born in Strasbourg, France on July 4, 1964. He received diplomas in Electrical Engineering from the Ecole Polytechnique, Palaiseau, France, and from Telecom University, Paris, in 1987 and 1989, respectively.

Since 1989, he has been with the Centre National d'Etudes des Télécommunications (CNET), Issy-Les-Moulineaux, France, where he is completing work in the Ph.D. degree in Signal Processing at Télécom University, specializing in wavelet theory, image coding, and fast signal algorithms.

Martin Vetterli was born in Switzerland in 1957. He received the Dip. El.-Ing. degree from the Eidgenössische Technische Hochschule Zürich, Switzerland, in 1981; the Master of Science degree from Stanford University, Stanford CA, in 1982; and the Doctorat ès Science Degree from the Ecole Polytechnique Fédérale de Lausanne, Switzerland, in 1986.

In 1982, he was a Research Assistant at Stanford University, and from 1983 to 1986 he was a researcher at the Ecole Polytechnique. He has worked for Siemens and AT&T Bell Laboratories. In 1986, he joined Columbia University in New York where he is currently associate professor of Electrical Engineering, member of the Center for Telecommunications Research, and codirector of the Image and Advanced Television Laboratory.

He is a senior member of the IEEE, a member of SIAM and ACM, a member of the MDSP committee of the IEEE Signal Processing Society, and of the editorial boards of *Signal Processing* and *Image Communication*. He received the Best Paper Award of EURASIP in 1984 for his paper on multidimensional subband coding, and the Research Prize of the Brown Boveri Corporation (Switzerland) in 1986 for his thesis. His research interests include multirate signal processing, wavelets, computational complexity, signal processing for telecommunications and digital video processing.

REFERENCES

[AKA90] A.N. Akansu, R.A. Haddad, and H. Caglar, "The Binomial QMF-Wavelet Transform for Multiresolution Signal Decomposition," submitted *IEEE Trans. Signal Proc.*, 1990.

[ALL77] J.B. Allen and L.R. Rabiner, "A Unified Approach to Short-Time Fourier Analysis and Synthesis," *Proc. IEEE*, Vol. 65, No. 11, pp. 1558-1564, 1977.

[ANI90] M. Antonini, M. Barlaud, P. Mathieu and I. Daubechies, "Image Coding Using Vector Quantization in the Wavelet Transform Domain," in *Proc. 1990 IEEE Int. Conf. Acoust., Speech, Signal Proc.*, Albuquerque, NM, Apr.3-6, 1990, pp. 2297-2300.

[AUS90] L. Auslander and I. Gertner, "Wide-Band Ambiguity Function and a.x+b group," in Signal Processing, Part I: Signal Processing Theory, L. Auslander, T. Kailath, S. Mitter eds., Institute for Mathematics and its Applications, Vol. 22, Springer Verlag, New York, pp.1-12, 1990.

[BAS89] M. Basseville and A. Benveniste, "Multiscale Statistical Signal Processing," in Proc. 1989 IEEE Int. Conf. Acoust., Speech, Signal Proc., Glasgow, Scotland, pp. 2065-2068, May 23-26, 1989.

[BER88] J. Bertrand and P. Bertrand, "Time-Frequency Representations of Broad-Band Signals," in *Proc. 1988 IEEE Int. Conf. Acoust., Speech, Signal Proc.*, New York, NY, Apr.11-14, 1988, pp.2196-2199.

[BEY89] G. Beylkin, R. Coifman and V. Rokhlin, "Fast Wavelet Transforms and Numerical Algorithms, I," submitted, Dec. 1989.

[BOU85] G. F. Boudreaux-Bartels, "Time-Varying Signal Processing Using the Wigner Distribution Time-Frequency Signal Representation," in *Adv. in Geophysical Data Proc.*, Vol. 2, pp. 33-79, Jai Press Inc., 1985.

[BUR83] P.J. Burt and E.H. Adelson, "The Laplacian Pyramid as a Compact Image Code," IEEE Trans. on Com., Vol. 31, No.4, pp. 532-540, April 1983.

[BUR89] P.J. Burt, "Multiresolution Techniques for Image Representation, Analysis, and 'Smart' Transmission," Proc. SPIE Conf. on Visual Communication and Image Processing., pp. 2-15, Philadelphia, PA, Nov. 1989.

[CAL64] A. Calderón, "Intermediate Spaces and Interpolation, the Complex Method," Studia Math., Vol. 24, pp. 113-190, 1964.

[CHO91] K.C. Chou, S. Golden, and A.S. Willsky, "Modeling and Estimation of Multiscale Stochastic Processes," in Proc. *1991 IEEE Int. Conf. Acoust., Speech, Signal Proc.*, Toronto, Ontario, Canada, pp. 1709-1712, May 14-17, 1991.

[CIP90] B.A. Cipra, "A New Wave in Applied Mathematics," *Science, Research News*, Vol. 249, August 24, 1990.

[CLA80] T.A.C.M. Classen and W.F.G. Mecklenbräuer, "The Wigner Distribution — A Tool for Time-Frequency Signal Analysis, Part I, II, III," *Philips J. Res.*, Vol.35, pp. 217-389, 1980.

[COH66] L. Cohen, "Generalized Phase-Space Distribution Functions," *J. Math. Phys.*, Vol. 7, No. 5, pp. 781-786, 1966.

[COH89] L. Cohen, "Time-Frequency Distribution - A Review," *Proc. IEEE*, Vol. 77, No. 7, pp. 941-981, 1989.

[COH90a] A. Cohen, I. Daubechies and J.C. Feauveau, "Biorthogonal Bases of Compactly Supported Wavelets," to appear in *Comm. Pure and Applied Math.*

[COH90b] A. Cohen, "Construction de Bases d'Ondelettes Hölderiennes," *Revista Matematica Iberoamericana*, Vol.6, No. 3 y 4, 1990.

[CRI76] A. Croisier, D. Esteban, and C. Galand, "Perfect Channel Splitting by Use of Interpolation, Decimation, Tree Decomposition Techniques," Int. Conf. on Information Sciences/Systems, Patras, pp. 443-446, Aug. 1976.

[CRO76] R.E. Crochiere, S.A. Weber, and J.L. Flanagan, "Digital Coding of Speech in Subbands," *Bell Syst. Tech. J.*, Vol.55, pp.1069-1085, Oct.1976.

[CRO83] R.E. Crochiere, and L.R. Rabiner, Multirate Digital Signal Processing, Prentice-Hall, Englewood Cliffs, NJ, 1983.

[DAU88] I. Daubechies, "Orthonormal Bases of Compactly Supported Wavelets," *Comm. in Pure and Applied Math.*, Vol.41, No.7, pp.909-996, 1988.

[DAU90a] I. Daubechies, "The Wavelet Transform, Time--Frequency Localization and Signal Analysis," *IEEE Trans. on Info. Theory*, Vol. 36, No.5, pp.961-1005, Sept. 1990.

[DAU90b] I. Daubechies, "Orthonormal Bases of Compactly Supported Wavelets II. Variations on a Theme," submitted to *SIAM J. Math. Anal.*

[DAU90c] I. Daubechies and J.C. Lagarias, "Two-Scale Difference Equations II. Local Regularity, Infinite Products of Matrices and Fractals," submitted to *SIAM J. Math. Anal*, 1990.

[DUF52] R.J. Duffin and A.C. Schaeffer, "A Class of Nonharmonic Fourier Series," *Trans. Am. Math. Soc.*, Vol. 72, pp. 341-366, 1952.

[EST77] D. Esteban and C. Galand, "Application of Quadrature Mirror Filters to Split-Band Voice Coding Schemes," Int. Conf. Acoust., Speech, Signal Proc., Hartford, Connecticut, pp. 191-195, May 1977.

[FLA89] P. Flandrin, "Some Aspects of Non-Stationary Signal Processing with Emphasis on Time-Frequency and Time-Scale Methods," in [WAV89], pp.68-98, 1989.

[FLA90] P. Flandrin and O. Rioul, "Wavelets and Affine Smoothing of the Wigner-Ville Distribution," in *Proc. 1990 IEEE Int. Conf. Acoust., Speech, Signal Proc.*, Albuquerque, NM, April 3-6, 1990, pp. 2455-2458.

[FOU88] J. B. J. Fourier, "Théorie Analytique de la Chaleur," in *Oeuvres de Fourier*, tome premier, G.Darboux, Ed., Paris: Gauthiers-Villars, 1888.

[FRA28] P. Franklin, "A Set of Continuous Orthogonal Functions," *Math. Annal.*, Vol.100, pp.522-529, 1928.

[FRA86] M. Frazier and B. Jawerth, "The φ-Transform and Decomposition of Distributions," *Proc. Conf. Function Spaces and Appl.*, Lund 1986, Lect. Notes Math., Springer.

[GAB46] D. Gabor, "Theory of Communication," *J. of the IEE*, Vol.93, pp.429-457, 1946.

[GAC91] N. Gache, P. Flandrin, and D. Garreau, "Fractal Dimension Estimators for Fractional Brownian Motions," in Proc. *1991 IEEE Int. Conf. Acoust., Speech, Signal Proc.*, Toronto, Ontario, Canada, pp. 3557-3560, May 14-17, 1991.

[GOU84] P. Goupillaud, A. Grossmann and J. Morlet, "Cycle-Octave and Related Transforms in Seismic Signal Analysis," *Geoexploration*, Vol.23, pp.85-102, Elsevier Science Pub., 1984/85.

[GRO84] A. Grossmann and J. Morlet, "Decomposition of Hardy Functions into Square Integrable Wavelets of Constant Shape," *SIAM J.Math.Anal.*, Vol.15, No.4, pp.723-736, July 1984.

[GRO89] A. Grossmann, R. Kronland-Martinet, and J. Morlet, "Reading and Understanding Continuous Wavelet Transforms," in [WAV89], pp.2-20, 1989.

[HAA10] A. Haar, "Zur Theorie der Orthogonalen Funktionensysteme," [in German] *Math. Annal.*, Vol. 69, pp. 331-371, 1910.

[HEI90] C.E. Heil, "Wavelets and Frames," in Signal Processing, Part I: Signal Processing Theory, L. Auslander, et al. eds., IMA, Vol. 22, Springer , New York, pp.147-160, 1990.

[HER71] O. Herrmann, "On the Approximation Probem in NonRecursive Digital Filter Design," *IEEE Trans. Circuit Theory*, Vol. CT-18, No. 3, pp. 411-413, May 1971.

[IT92] *IEEE Transactions on Information Theory*, Special issue on Wavelet Transforms and Multiresolution Signal Analysis, to appear January, 1992.

[JOH80] J.D. Johnston, "A Filter Family Designed for Use in Quadrature Mirror Filter Banks," in *Proc. 1980 IEEE Int. Conf. Acoust., Speech, Signal Proc.*, pp. 291-294, Apr. 1980.

[KAD91] S. Kadambe and G.F. Boudreaux-Bartels, "A Comparison of the Existence of 'Cross Terms' in the Wavelet Transform, Wigner Distribution and Short-Time Fourier Transform," Submitted *IEEE Trans. Signal Proc.*, Revised Jan. 1991.

[KOV92] J. Kovacevic and M. Vetterli, "Non-separable Multidimensional Perfect Reconstruction Filter Banks and Wavelet Bases for R^n," *IEEE Trans. on Info. Theory*, Special Issue on wavelet transforms and multiresolution signal analysis, to appear Jan. 1992.

[LIT37] J. Littlewood and R. Paley, "Theorems on Fourier Series and Power Series," *Proc. London Math. Soc.*, Vol.42, pp. 52-89, 1937.

[MAL89a] S. Mallat, "A Theory for Multiresolution Signal Decomposition: the Wavelet Representation," *IEEE Trans. on Pattern Analysis and Machine Intell*, Vol.11, No. 7, pp.674-693, July 1989.

[MAL89b] S. Mallat, "Multifrequency Channel Decompositions of Images and Wavelet Models," *IEEE Trans. Acoust., Speech, Signal Proc.*, Vol. 37, No.12, pp.2091-2110, December 1989.

[MAL89c] S. Mallat, "Multiresolution Approximations and Wavelet Orthonormal Bases of $L^2(R)$,", *Trans. Amer. Math. Soc.*, Vol.315, No.1, pp.69-87, September 1989.

[MAL89d] S. Mallat and S. Zhong, "Complete Signal Representation with Multiscale Edges," submitted to *IEEE Trans. Pattern Analysis and Machine Intell.* 1989.

[MEY89] Y. Meyer, "Orthonormal Wavelets," in [WAV89], pp. 21-37, 1989.

[MEY90] Y. Meyer, *Ondelettes et Opérateurs*, Tome I. *Ondelettes*, Herrmann ed., Paris, 1990.

[MIN85] F. Mintzer, "Filters for Distortion-Free Two-Band Multirate Filter Banks," *IEEE Trans. on Acoust., Speech, Signal Proc.*, Vol.33, pp.626-630, June 1985.

[POR80] M.R. Portnoff, "Time-Frequency Representation of Digital Signals and Systems Based on Short-Time Fourier Analysis," *IEEE Trans. on Acoust., Speech, Signal Proc.*, Vol.28, pp.55-69, Feb. 1980.

[RAM88] T.A. Ramstad and T. Saramäki, "Efficient Multirate Realization for Narrow Transition-Band FIR Filters," in *Proc. 1988 IEEE Int. Symp. Circuits and Systems*, Helsinki, Finland, pp. 2019-2022, 1988.

[RIO90a] O. Rioul and P. Flandrin, "Time-Scale Energy Distributions: A General Class Extending Wavelet Transforms," to appear in *IEEE Trans. Signal Proc.*

[RIO90b] O. Rioul, "A Discrete-Time Multiresolution Theory Unifying Octave-Band Filter Banks, Pyramid and Wavelet Transforms," submitted to *IEEE Trans. Signal Proc.*

[RIO91a] O. Rioul and P. Duhamel, "Fast Algorithms for Discrete and Continuous Wavelet Transforms," submitted to *IEEE Trans. Information Theory*, Special Issue on Wavelet Transforms and Multiresolution Signal Analysis.

[RIO91b] O. Rioul, "Dyadic Up-Scaling Schemes: Simple Criteria for Regularity," submitted to *SIAM J. Math. Anal.*

[ROS84] A. Rosenfeld ed., Multiresolution Techniques in Computer Vision, Springer-Verlag, New York 1984.

[SMI86] M.J.T. Smith and T.P. Barnwell, "Exact Reconstruction for Tree-Structured Subband Coders," *IEEE Trans. on Acoust., Speech and Signal Proc.*, Vol. ASSP-34, pp. 434-441, June 1986.

[SHE90] M.J. Shensa, "The Discrete Wavelet Transform: Wedding the à Trous and Mallat Algorithms," submitted to *IEEE Trans. on Acoust., Speech, and Signal Proc.*, 1990.

[SPE67] J.M. Speiser, "Wide-Band Ambiguity Functions," *IEEE Trans. on Info. Theory*, pp. 122-123, 1967.

[VAI87] P.P. Vaidyanathan, "Quadrature Mirror Filter Banks, M-band Extensions and Perfect-Reconstruction Techniques," *IEEE ASSP Magazine*, Vol. 4, No. 3, pp.4-20, July 1987.

[VAI88] P.P. Vaidyanathan and P.-Q. Hoang, "Lattice Structures for Optimal Design and Robust Implementation of Two-Band Perfect Reconstruction QMF Banks," *IEEE Trans. on Acoust., Speech and Signal Proc.*, Vol. ASSP-36, No. 1, pp.81-94, Jan. 1988.

[VAI89] P.P. Vaidyanathan and Z. Doganata, "The Role of Lossless Systems in Modern Digital Signal Processing," *IEEE Trans. Education*, Special issue on Circuits and Systems, Vol. 32, No.3, Aug. 1989, pp.181-197.

[VET86] M. Vetterli, "Filter Banks Allowing Perfect Reconstruction," *Signal Processing*, Vol.10, No.3, April 1986, pp.219-244.

[VET89] M. Vetterli and D. Le Gall, "Perfect Reconstruction FIR Filter Banks: Some Properties and Factorizations, *IEEE Trans. on Acoust., Speech Signal Proc.*, Vol.37, No.7, pp.1057-1071, July 1989.

[VET90a] M. Vetterli and C. Herley "Wavelets and Filter Banks: Relationships and New Results," in *Proc. 1990 IEEE Int. Conf. Acoust., Speech, Signal Proc.*, Albuquerque, NM, pp. 1723-1726, Apr. 3-6, 1990.

[VET90b] M. Vetterli and C. Herley, "Wavelets and Filter Banks: Theory and Design," to appear in *IEEE Trans. on Signal Proc.*, 1992.

[WAV89] *Wavelets, Time-Frequency Methods and Phase Space*, Proc. Int. Conf. Marseille, France, Dec. 14-18, 1987, J.M. Combes et al. eds., Inverse Problems and Theoretical Imaging, Springer, 315 pp., 1989.

[WIC89] M.V. Wickerhauser, "Acoustic Signal Compression with Wave Packets," preprint Yale University, 1989.

[WOO53] P.M. Woodward, *Probability and Information Theory with Application to Radar*, Pergamon Press, London, 1953.

[WOR90] G.W. Wornell, "A Karhunen-Loève-like Expansion for 1/f Processes via Wavelets," *IEEE Trans. Info. Theory*, Vol. 36, No. 4, pp.859-861, July 1990.

[YOU78] J.E. Younberg, S.F. Boll, "Constant-Q Signal Analysis and Synthesis," in Proc. 1978 IEEE Int. Conf. on Acoust., Speech, and Signal Proc., Tulsa, OK, pp. 375-378, 1978.

EXTENDED REFERENCES

History of Wavelets: see [HAA10, FRA28, LIT37, CAL64, YOU78, GOU84].

Books on Wavelets: (see also [WAV89, MEY90])

I. Daubechies, *Ten Lectures on Wavelets*, CBMS, SIAM publ., to appear.

Wavelets and their Applications, R.R. Coifman, I. Daubechies, S. Mallat, Y. Meyer scientific eds., L.A. Raphael, M.B. Ruskai managing eds., Jones and Bertlel pub., to appear, 1991.

Tutorials on Wavelets: (see also [FLA89, GRO89, MAL89b, MEY89])

R.R. Coifman, "Wavelet Analysis and Signal Processing," in Signal Processing, Part I: Signal Processing Theory, L. Auslander et al. eds., IMA, Vol. 22, Springer, New York, 1990.

C.E. Heil and D.F. Walnut, "Continuous and Discrete Wavelet Transforms," *SIAM Review*, Vol. 31, No. 4, pp 628-666, Dec. 1989.

Y. Meyer, S. Jaffard, O. Rioul, "L'Analyse par Ondelettes," [in French]*Pour La Science*, No.119, pp.28-37, Sept 1987.

G. Strang, "Wavelets and Dilation Equations: A Brief Introduction," *SIAM Review*, Vol. 31, No. 4, pp. 614-627, Dec. 1989.

Mathematics, Mathematical Physics and Quantum Mechanics: (see also [GRO84, DAU88, MEY90])

G. Battle, "A Block Spin Construction of Ondelettes, II. The Quantum Field Theory (QFT) Connection," Comm. Math. Phys., Vol. 114, pp. 93-102, 1988.

W.M. Lawton, "Necessary and Sufficient Conditions for Constructing Orthonormal Wavelet Bases," Aware Tech. Report # AD900402.

P.G. Lemarié and Y. Meyer, "Ondelettes et Bases Hilbertiennes," [in French] *Revista Matematica Iberoamericana*, Vol.2, No.1&2, pp.1-18, 1986.

T. Paul, "Affine Coherent States and the Radial Schrödinger Equation I. Radial Harmonic Oscillator and the Hydrogen Atom," to appear in *Ann. Inst. H.Poincaré*.

H.L. Resnikoff, "Foundations or Arithmeticum Analysis: Compactly Supported Wavelets and the Wavelet Group," Aware Tech. Report # AD890507.

Regular Wavelets: see [DAU88, DAU90b, DAU90c, COH90b, RIO91b]

Computer-Aided Geometric Design using Regular Interpolators:

S.Dubuc, "Interpolation Through an Iterative Scheme," *J. Math. Analysis Appl.*, Vol.114, pp.185-204, 1986.

N. Dyn and D. Levin, "Uniform Subdivision Schemes for the Generation of Curves and Surfaces," Constructive Approximation, to appear.

Numerical Analysis: (see also [BEY89])

R.R. Coifman, "Multiresolution Analysis in Nonhomogeneous Media," in [WAV89], pp. 259-262, 1989.

V. Perrier, "Toward a Method to Solve Partial Differential Equations Using Wavelet Bases," in [WAV89], pp. 269-283, 1989.

Multiscale Statistical Signal Processing: see [BAS89, CHO91].

Fractals, Turbulence: (see also [GAC91, WOR90])

A. Arnéodo, G. Grasseau, and M. Holschneider, "Wavelet Transform of Multifractals," *Phys. Review Letters*, Vol.61, No.20, pp.2281-2284, 1988.

F. Argoul, A. Arnéodo, G. Grasseau, Y. Gagne, E.J. Hopfinger, and U. Frisch, "Wavelet Analysis of Turbulence Reveals the Multifractal Nature of the Richardson Cascade," *Nature*, Vol.338, pp.51-53, March 1989.

One-Dimensional Signal Analysis: (see also [GRO89], [WIC-89])

C. D'Alessandro and J.S. Lienard, "Decomposition of the Speech Signal into Short-Time Waveforms Using Spectral Segmentation," in *Proc. 1988 IEEE Int. Conf. Acoust., Speech, Signal Proc.*, New York, Apr.11-14, 1988, pp.351-354.

S. Kadambe and G.F. Boudreaux-Bartels, "A Comparison of Wavelet Functions for Pitch Detection of Speech Signals," in *Proc. 1991 IEEE Int. Conf. Acoust., Speech, Signal Proc.*, Toronto, Ontario, Canada, pp. 449-452, May. 14-17, 1991.

R. Kronland-Martinet, J. Morlet, and A. Grossmann, "Analysis of Sound Patterns Through Wavelet Transforms," *Int. J. Pattern Recognition and Artificial Intelligence*, Vol.1, No.2, pp. 273-302, pp.97-126, 1987.

J.L. Larsonneur and J. Morlet, "Wavelets and Seismic Interpretation," in [WAV89], pp.126-131, 1989.

F. B. Tuteur, "Wavelet Transformations in Signal Detection," in *Proc. 1988 IEEE Int. Conf. Acoust., Speech, Signal Proc.*, New York, NY, Apr. 11-14, 1988, pp.1435-1438. Also in [WAV89], pp. 132-138, 1989.

Radar/Sonar, Ambiguity Functions: see e.g. [AUS90, SPE67, WOO53]

Time-Scale Representations: see [BER88, FLA89, FLA90, RIO90a].

Filter Bank Theory: (see also [CRI76, CRO76, EST77, CRO83, MIN85, SMI86, VAI87, VAI88, VAI89, VET86, VET89])

J.D. Johnston, "A Filter Family Designed for Use in Quadrature Mirror Filter Banks," *Proc. ICASSP-80*, pp.291-294, April 1980.

T.Q. Nguyen and P.P. Vaidyanathan, "Two-Channel Perfect--Reconstruction FIR QMF Structures Which Yield Linear-Phase Analysis and Synthesis Filters," *IEEE Trans. Acoust., Speech, Signal Processing*, Vol. ASSP-37, No. 5, pp.676-690, May 1989.

M.J.T. Smith and T.P. Barnwell, "A New Filter Bank Theory for Time-Frequency Representation," *IEEE Trans. on Acoust., Speech and Signal Proc.*, Vol. ASSP-35, No.3, March 1987, pp. 314-327.

P.P. Vaidyanathan, *Multirate Filter Banks*, Prentice Hall, to appear.

Pyramid Transforms: see [BUR83, BUR89, ROS84].

Multidimensional Filter Banks: (see also [KOV92])

G. Karlsson and M. Vetterli, "Theory of Two-Dimensional Multirate Filter Banks," *IEEE Trans. on Acoust., Speech, Signal Proc.*, Vol.38, No.6, pp.925-937, June 1990.

M. Vetterli, "Multi-Dimensional Subband Coding: Some Theory and Algorithms," *Signal Processing*, Vol. 6, No.2, pp. 97-112, Feb. 1984.

M. Vetterli, J. Kovacevic and D. Le Gall, "Perfect Reconstruction Filter Banks for HDTV Representation and Coding," *Image Communication*, Vol.2, No.3, Oct.1990, pp.349-364.

E. Viscito and J. Allebach, "The Analysis and Design of Multidimensional FIR Perfect Reconstruction Filter Banks for Arbitrary Sampling Lattices," *IEEE Trans. Circuits and Systems*, Vol.38, pp.29-42, Jan. 1991.

Multidimensional Wavelets: (see also [ANI90, KOV92, MAL89a, MAL89b, MAL89c, MAL89d]).

J.C. Feauveau, "Analyse Multirésolution pour les Images avec un Facteur de Résolution 2," [in French], Traitement du Signal, Vol. 7, No. 2, pp. 117-128, July 1990.

K. Gröchenig and W.R. Madych, "Multiresolution analysis, Haar bases and self-similar tilings of R^n," submitted to *IEEE Trans. on Info. Theory*, Special Issue on wavelet transforms and multiresolution signal analysis, Jan.1992.