A Theory for Multiresolution Signal Decomposition: The Wavelet Representation

Stephane G. Mallat
University of Pennsylvania

A THEORY FOR MULTIRESOLUTION SIGNAL DECOMPOSITION: THE WAVELET REPRESENTATION

S. G. Mallat

MS-CIS-87-22
GRASP LAB 103

Department Of Computer and Information Science
School of Engineering and Applied Science
University of Pennsylvania
Philadelphia, PA 19104-6389

May 1987

Acknowledgements: This work is supported in part NSF-CER/DCR82-19196 AO2, NSF-CER MCS-8219196, NSF/DCR-8410771, Air Force/F49620-85-K-0018, ARMY DAAG-29-84-K-0061, DAA29-84-9-0027 and DARPA NOO14-85-K-0807 P002.

A THEORY FOR MULTIRESOLUTION SIGNAL DECOMPOSITION : THE WAVELET REPRESENTATION

STEPHANE G. MALLAT

GRASP lab, Dept of Computer and Information Science
University of Pennsylvania Philadelphia, PA 19104-6389
Net address: mallat@grasp.cis.upenn.edu.arpa

ABSTRACT

It is now well admitted in the computer vision literature that a multiresolution decomposition provides a useful image representation for vision algorithms. In this paper we show that the *wavelet theory* recently developed by the mathematician Y. Meyer enables us to understand and model the concepts of resolution and scale. In computer vision we generally do not want to analyze the images at each resolution level because the information is redundant. After processing the signal at a resolution r_0 , it is more efficient to analyze only the *additional details* which are available at a higher resolution r_1 . We prove that this difference of information can be computed by decomposing the signal on a wavelet orthonormal basis and that it can be efficiently calculated with a pyramid transform. This can also be interpreted as a division of the signal in a set of orientation selective frequency channels. Such a decomposition is particularly well adapted for computer vision applications such as signal coding , texture discrimination , edge detection , matching algorithms and fractal analysis.

1. Introduction

In the physical world, the important structures to recognize are of very different sizes. Furthermore, depending on the distance to the focal plane of a video camera, these elements will appear at different scales on the image. In computer vision we would like to analyze each structure and at the same time process the minimum amount of details necessary to recognize them. One strategy is to process gradually the details with an increasing resolution until the recognition is achieved. For this purpose Witkin [28] has created the scale space representation in which the image appears at several resolutions (or scales) $(r_j)_{0 \leq j \leq N}$ $(r_j < r_{j+1})$. With this representation, after analyzing the image at a given resolution r_j , for more details we have to process the image at a higher resolution r_{j+1} . The images at resolutions r_j and r_{j+1} give a lot of redundant information. It would be more efficient to process only the *additional* details available at the resolution r_{j+1} , which we shall call the *detail signal* at resolution r_{j+1} . For this purpose we will define a new *discrete* representation called the *wavelet representation* which provides a coarse approximation of the signal (resolution r_0) plus the detail signals at the successive resolutions r_j for $1 \leq j < N$.

The concepts of scale and resolution are interdependent in computer vision. The resolution of an object image gives the minimum size of the object details which will be distinguishable in the image. It is a function of the number of photoreceptors which have measured the light reflected by the object. As there is a finite number of photoreceptors per unit area in the camera focal plane, the resolution will thus depends

This work is supported in part by NSF-CER/DCR82-19196 A02, NSF/DCR-8410771, Air Force/F49620-85-K-0018, ARMY DAAG-29-84-K-0061, and DARPA/ONR N0014-85-K-0807.

on the size (or scale) of the projection of the object on this plane. In this paper we will first model the scaling transform which modifies the resolution of the image components. This will then enable us to mathematically define the detail signal at each resolution through a new interpretation of the wavelet theory. Wavelets are particular functions studied by the mathematician Y. Meyer [19] which allow us to build interesting orthonormal bases of $\mathbf{L}^2(\mathbf{R}^n)$. We will show that the detail signal can be computed by decomposing the original signal on such an orthonormal basis and that it can be efficiently calculated with a pyramid architecture using quadrature mirror filters. The wavelet representation is complete and we will give a similar algorithm for reconstructing the original signal from its decomposition. Like Marr's [18] model the wavelet representation can be interpreted as a decomposition of the original signal into a set of independent orientation selective frequency channels. For two-dimensional signals such as images we will present a separable decomposition which privileges the horizontal and vertical orientations. However the mathematical model enables us to build non-separable representations with as many orientation tunings as desired.

The wavelet representation has many applications in computer vision and signal processing in general. We will describe in particular how it can be used for signal matching , data compression , edge detection , texture discrimination and fractal analysis. Finally, the wavelet representation will be compared with the DOG and Gabor representations currently used in computer vision and having some similar features. We have tried in this paper to give the mathematical insights of the model without justifying in details all the results. The main steps of the theorem proofs are given in the appendices.

2. Scaling transform

We shall hereunder define a scaling transform for one-dimensional signals but the definitions and theorems can be extended to any dimension. We will then study more precisely the two-dimensional case for image processing applications.

2.1. Characterization and modeling

Let T_r be the scaling transform which associates to any signal an approximation at a resolution r . Intuitively the resolution r gives the minimum size of the details which can still be found in the approximation, r being defined with respect to a certain distance unit. As described in the introduction, we want to compute the approximation of the original signal at different resolutions and then calculate the detail signals. Koenderink [13] has shown that in order to obtain consecutive detail signals carrying approximatively a constant amount of information, the sequence of resolutions should vary exponentially : $(r^j)_{j\in\mathbf{Z}}$ $(r > 1)$. In order to simplify the computer implementation we will choose r = 2 , but the model can be extended for any $r \in \mathbf{Q}$.

We will suppose that our original signal f(x) has a finite energy :

$$\int_{-\infty}^{+\infty} |f(x)|^2 dx < +\infty \quad <==> \quad f(x) \in \mathbf{L}^2(\mathbf{R}).$$

It is well known that there exists an inner product in the vector space $\mathbf{L}^2(\mathbf{R})$.

$$\forall f(x) \in \mathbf{L}^2(\mathbf{R}) \ , \forall g(x) \in \mathbf{L}^2(\mathbf{R}) \qquad < f \ , g > = \int_{-\infty}^{+\infty} f(x)\, g(x)\, dx \ .$$

In the computer vision literature certain authors such as Yuille & Poggio [30] and Koenderink [13] have defined several principles forcharacterizing a scaling transform. Hereunder we will describe six such principles with their mathematical interpretation and build a global model characterizing a scaling transform. Having selected the sequence of resolution levels $(2^j)_{j\in\mathbf{Z}}$, we will restrict our analysis to the scaling transforms T_{2^j} which approximate a signal at a resolution 2^j .

- T_{2^j} is a linear transformation and if g(x) is the approximation of a signal at the resolution 2^j than g(x) is invariant by T_{2^j} .

T_{2^j} can thus be interpreted as a projection operator on a particular vector space V_j included in $\mathbf{L}^2(\mathbf{R})$. V_j is the set of all possible approximated signals at the resolution 2^j . We will suppose that the projection

is orthogonal (we can always find an inner product for which this is true).

- Causality: The approximation of a signal at resolution 2^{j+1} contains all the necessary informations to build the same signal at a smaller resolution 2^j .

Since T_{2^j} is a projection operator on V_j this principle is equivalent to :

$$\forall j \in \mathbf{Z}, \quad V_j \subset V_{j+1} . \tag{1}$$

- A scaling transform does not *privilege* a priori any particular resolution level.

The approximations at different resolutions should thus be derived from one another by using a constant scaling factor. For the sequence of resolutions $(2^j)_{j\in\mathbf{Z}}$ this scaling factor is 2 .

$$\forall j \in \mathbf{Z}, \quad g(x) \in V_j \quad \Longleftrightarrow \quad g(2x) \in V_{j+1} . \tag{2}$$

- An approximation of a signal at a resolution 2^j can be characterized by 2^j samples per unit of length.

Because of the previous principle this assertion is valid for all $j \in \mathbf{Z}$ if it is true for j= 0 . It can thus be written

$$\exists I_0 \quad \text{isomorphism from} \quad V_0 \quad \text{onto} \quad l^2(\mathbf{Z}) \tag{3}$$

$$\text{where} \quad l^2(\mathbf{Z}) = \{(\alpha_i)_{i\in\mathbf{Z}} \ / \ \sum_{i=-\infty}^{+\infty} |\alpha_i|^2 < +\infty\}$$

- When a function is translated by a length proportional to 2^j , its approximation at the resolution 2^j is translated by the same length and it will be characterized by the same samples which are translated as well.

Because of (2) it is sufficient to express the above principle for j = 0 .

$$\forall f(x) \in \mathbf{L}^2(\mathbf{R}) , \forall k \in \mathbf{Z} \quad T_0(f(x)) = g(x) \quad \Longleftrightarrow \quad T_0(f(x-k)) = g(x-k) \tag{4}$$

$$\forall k \in \mathbf{Z} , \forall f(x) \in V_0 \quad I_0(g(x)) = (\alpha_i)_{i\in\mathbf{Z}} \quad \Longleftrightarrow \quad I_0(g(x-k)) = (\alpha_{i-k})_{i\in\mathbf{Z}} . \tag{5}$$

- When the resolution increases towards $+\infty$ the approximated signal converges towards the original signal. Conversely when the resolution decreases to zero, the approximated signal contains less and less information and should ultimately converge towards zero.

Since the approximated signal at a resolution 2^j is equal to the orthogonal projection on a space V_j , when j increases the space V_j should ultimately cover almost all the initial vector space $\mathbf{L}^2(\mathbf{R})$. When j decreases, it should shrink up to the vector space $\{0\}$. Property (1) enables us to write this :

$$\bigcup_{j=-\infty}^{j=+\infty} V_j \quad \text{is dense in} \quad \mathbf{L}^2(\mathbf{R}) \qquad \text{and} \qquad \bigcap_{j=-\infty}^{j=+\infty} V_j = \{0\} . \tag{6}$$

We will call any set of vector spaces $\left[V_j\right]_{j\in\mathbf{Z}}$ which verifies the properties (1) to (6) a *multiresolution vector space sequence.*

From these six principles, we have shown that a scaling transform at a resolution 2^j is an orthonormal projection on a space V_j of a multiresolution vector space sequence. In order to compute this orthonormal projection for any signal f(x) , we need to find an orthonormal basis of V_j . The following theorem shows that such an orthonormal basis can be computed from a unique function which totally characterizes the multiresolution vector space sequence.

Theorem 1

Let $\left(V_j\right)_{j \in \mathbf{Z}}$ be a multiresolution space sequence, then there exists a unique function $\phi(x)$ called a *scaling function* such that for any $j \in \mathbf{Z}$ if $\phi^j(x) = \sqrt{2^j}\,\phi(2^j x)$ then

$$\left(\phi^j(x - 2^{-j}n)\right)_{n \in \mathbf{Z}} \quad \text{is an orthonormal basis of} \quad V_j \ .$$

Appendix 1 gives the intermediate steps of the proof. We can therefore build an orthonormal basis of any V_j by scaling the function $\phi(x)$ with a coefficient 2^j and translating the resulting function on a grid whose interval is proportional to 2^{-j} . The orthogonal projection on V_j can now be computed by decomposing the signal f(x) on this orthonormal basis.

Let P_{V_j} be the orthogonal projection operator on the vector space V_j and let $\phi_k^j(x) = \phi^j(x - 2^{-j}n)$,

$$\text{then} \quad T_{2^j}(f)(x) = P_{V_j}(f)(x) = \sum_{n=-\infty}^{+\infty} < f \ , \phi_k^j > \ \phi_k^j(x) \tag{7}$$

The approximation of the signal f(x) at the resolution 2^j is thus characterized by the set of samples :

$$\left(< f \ , \phi_k^j > \right)_{n \in \mathbf{Z}} .$$

As an example we will compute these coefficients for a function f(x) which is constant on a large interval. $f(x) = c$ for $x \in [2^{-j}n - N \ , 2^{-j}n + M]$ with $N \ , M >> 2^{-j}$ and $f(x) = 0$ outside this interval.

$$< f \ , \phi_k^j > = c \int_{2^{-j}n-N}^{2^{-j}n+N} \phi_k^j(x)\,dx \ = \frac{c}{\sqrt{2^j}} \int_{-2^jN}^{2^jM} \phi(x)\,dx \ \approx \ \frac{c}{\sqrt{2^j}} \int_{-\infty}^{+\infty} \phi(x)\,dx$$

We can show (appendix 2 (45)) that

$$\int_{-\infty}^{+\infty} \phi(x)\,dx \ = 1 \qquad \text{so} \quad < f \ , \phi_k^j > \approx \frac{c}{\sqrt{2^j}} \ .$$

For such a function, to have coefficients of same value at all resolutions we need to multiply them by $\sqrt{2^j}$. For any signal f(x) , the set of samples

$$S_j = \left(\sqrt{2^j} < f \ , \phi_k^j > \right)_{n \in \mathbf{Z}} . \tag{8}$$

will be called a *discrete approximation* of f(x) at the resolution 2^j . Since computers can only process discrete signals we will concentrate on these discrete approximations, which completely characterize the continuous approximations. Since

$$< f \ , \phi_k^j > \ = \int_{-\infty}^{+\infty} f(x)\,\phi_k^j(x)\,dx \ = \int_{-\infty}^{+\infty} f(x)\,\phi^j(x - 2^{-j}n)\,dx \ = \ (f(x) * \phi^j(-x))(2^{-j}n) \ ,$$

S_j is thus also equal to the convolution of f(x) with $\phi^j(-x)$, sampled at the rate 2^j .

$$S_j = \left(\sqrt{2^j}\ (f(x) * \phi^j(-x))(2^{-j}n) \right)_{n \in \mathbf{Z}} . \tag{9}$$

As predicted by the fourth principle, we are now able to characterize the approximation of a signal at a resolution 2^j with a discrete signal of 2^j samples per length unit.

Fig. 1 shows an exponentially decreasing scaling function. Its Fourier transform has the shape of a low-pass filter, hence equation (8) can be interpreted as a low-pass filtering of the signal. Indeed, we remove all the details smaller than 2^j which correspond to the higher frequencies of the signal. We are now going to study a practical procedure to compute the discrete approximations at different resolutions.

2.2. Implementation of a scaling transform

In computer vision as often in signal processing we can only process discrete signals. The measuring device low-passes the input continuous signal and the digitizer provides a uniform sampling at the output. The distance unit will be taken equal to the sampling interval. We can thus consider that these samples give a discrete approximation of the original signal at the resolution 1 : S_0 .

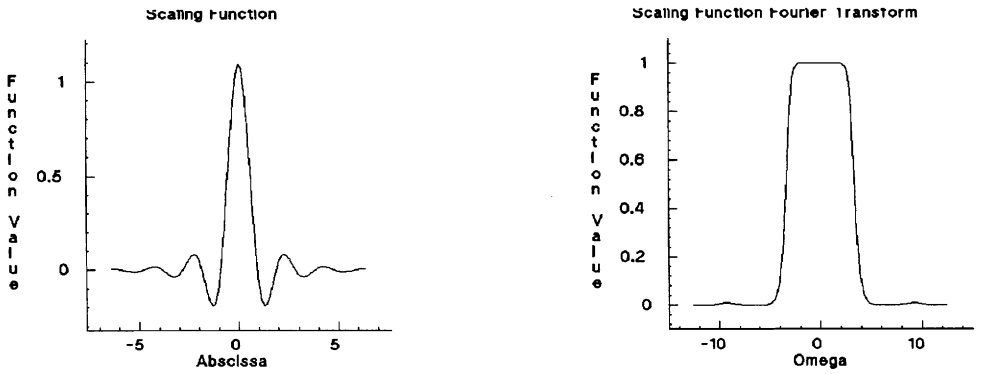


Fig. 1. *Example of scaling function with its Fourier transform. It decreases exponentially in the spatial domain and like* $\frac{1}{\omega^4}$ *in the frequency domain.*

The causality principle shows that from S_0 we can compute all the S_j for j < 0 . We shall now prove that this can be performed with a simple recursive algorithm.

For $\phi(x)$ a scaling function, $\left[\phi_k^{j+1}(x)\right]_{j \in Z}$ is an orthonormal basis of V_{j+1} so $\phi_k^j(x) \in V_j \subset V_{j+1}$ can be decomposed in this basis:

$$\phi_k^j(x) = \sum_{k=-\infty}^{+\infty} < \phi_k^j , \phi_l^{j+1} > \phi_l^{j+1}(x) . \tag{10}$$

$$< \phi_k^j , \phi_l^{j+1} > = \int_{-\infty}^{+\infty} \sqrt{2^j}\, \phi(2^j x - n)\, \sqrt{2^{j+1}}\, \phi(2^{j+1}x - k)\, dx = \int_{-\infty}^{+\infty} \sqrt{2^{-1}}\, \phi(2^{-1}x)\ \phi(x-(k-2n))\, dx$$

$$\text{thus,} \quad < \phi_k^j , \phi_l^{j+1} > = < \phi_0^{-1} , \phi_{k-2n}^0 > \tag{11}$$

so by writing the inner product of f(x) with the functions of both sides of equation (9) and multiplying by $\sqrt{2^{j+1}}$, we have

$$\sqrt{2^j} < f , \phi_k^j > = \sum_{k=-\infty}^{+\infty} \sqrt{2^{j+1}} < f , \phi_l^{j+1} > \frac{1}{\sqrt{2}} < \phi_0^{-1} , \phi_{k-2n}^0 > .$$

Let H be a discrete filter with impulse response

$$h(n) = \frac{1}{\sqrt{2}} < \phi_0^{-1} , \phi_n^0 > \tag{12}$$

and $\tilde{H}$ be the symmetrical filter with impulse response $\tilde{h}(n) = h(-n)$,

$$\sqrt{2^j} < f , \phi_k^j > = \sum_{k=-\infty}^{+\infty} \sqrt{2^{j+1}} < f , \phi_l^{j+1} > \tilde{h}(2n-k) . \tag{13}$$

Equation (13) shows that S_j can be computed by convolving S_{j+1} with $\tilde{H}$ and by keeping every other sample of the output . All the S_j for j < 0 can thus be computed from S_0 by repeating this process for all $j < 0$. This is called a pyramid transform. In practice the measuring device only gives a finite number N of samples. We can then easily show that each discrete signal S_j (j < 0) will have $2^j N$ samples. Fig. 2 shows the discrete approximated signal S_j of a continuous signal f(x) , for j = 0 , -1 , -2 , -3 . We have approximated the impulse response of the filter H shown in Fig. 3 by taking h(n) = 0 for |n| > 8 . The continuous approximated signals $T_{2^j}(f)$ have been calculated with equation (7) . When the resolution decreases the smaller details of f(x) gradually disappear.

Let us insert equations (11) and (12) in (10) . For j = -1 and n = 0 we have :

$$\frac{1}{2}\, \phi(\frac{x}{2}) = \sum_{k=-\infty}^{+\infty} h(k)\, \phi(x-k) \tag{14}$$

Let $\hat{\phi}(\omega)$ and $H(\omega)$ be respectively the Fourier transform of $\phi(x)$ and the discrete Fourier transform of h(n). The Fourier transform of equation (14) can be written

$$\hat{\phi}(2\omega) = H(\omega)\ \hat{\phi}(\omega) \tag{15}$$

This equation gives a simple relation between the filter H and the corresponding scaling function $\phi(x)$. Based on this equation, the following theorem gives a complete characterization of the filter H and shows that given such a filter it is possible to build $\phi(x)$. This will enable us to numerically compute some scaling functions. In computer vision we often need to analyze the signal locally both in the spatial and frequency domains. For this purpose, we will look for scaling functions which are well localized in both domains. In particular we will suppose that $\phi(x) = O(\frac{1}{x^3})$ in order to simplify the statement of the theorem.

Theorem 2

Let $\phi(x)$ be a scaling function and H a discrete filter with impulse response $h(n) = \frac{1}{\sqrt{2}} < \phi_0^{-1}, \phi_n^0 >$ and discrete Fourier transform $H(\omega)$.

H is a low pass filter with the following properties :

(a) $H(\omega)$ is 2π periodic, differentiable and $|H(0)| = 1$.

(b) $|H(\omega)|^2 + |H(\omega+\pi)|^2 = 1$

Conversely let H be a discrete filter with Fourier transform $H(\omega)$ satisfying (a), (b) and such that

(c) $|H(\omega)| \neq 0$ for $\omega \in [0, \pi/2]$

$$\text{then,} \quad \hat{\phi}(\omega) = \prod_{p=1}^{+\infty} H(2^{-p}\omega) \tag{16}$$

is the Fourier transform of a scaling function, and

$$\frac{1}{\sqrt{2}} < \phi_0^{-1}, \phi_n^0 > = h(n)$$

Appendix 2 gives some indications to prove this theorem. The filters which verify the property (b) are called *quadrature* filters. We can find an extensive description of such filters and numerical methods to synthesize them in the signal processing literature [5,20,25]. Given a quadrature filter H which verifies (a), (b) and (c) we can then compute the corresponding scaling function with equation (16). It is possible to choose the filter H so that the scaling function $\phi(x)$ will be well localized in both the frequency and spatial domains. In appendix 3 we give a class of symmetrical scaling functions which are exponentially decreasing and whose Fourier transforms decrease like $\frac{1}{\omega^n}$ for $n \in \mathbf{N}$. The scaling function shown in Fig. 1 is one of them. We can also find quadrature filters H having a Finite Impulse Response ($h(n) = 0$ for $|n| > n_0$) [26]. This leads to non symmetrical scaling functions which have a compact support [3] in the spatial domain. Fig. 3 shows the H filter of the scaling function given in Fig. 1.

3. The wavelet representation

3.1. Detail signal modeling

We do not want to process the signal at each resolution because the information is redundant. It is more efficient to successively process the details which can be found at the resolution 2^{j+1} but not at the resolution 2^j. The approximations of the signal at the resolution 2^{j+1} and 2^j are given respectively by the orthogonal projection of the signal on V_{j+1} and V_j. By applying the projection theorem we can easily show that the detail signal at the resolution 2^{j+1} are given by the orthogonal projection of the original signal on a vector space O_j such that:

$$O_j \quad \text{is orthogonal to} \quad V_j \tag{17}$$

$$O_j \oplus V_j = V_{j+1}\ . \tag{18}$$

For computing this orthogonal projection for any signal f(x) we need to find an orthonormal basis of O_j.

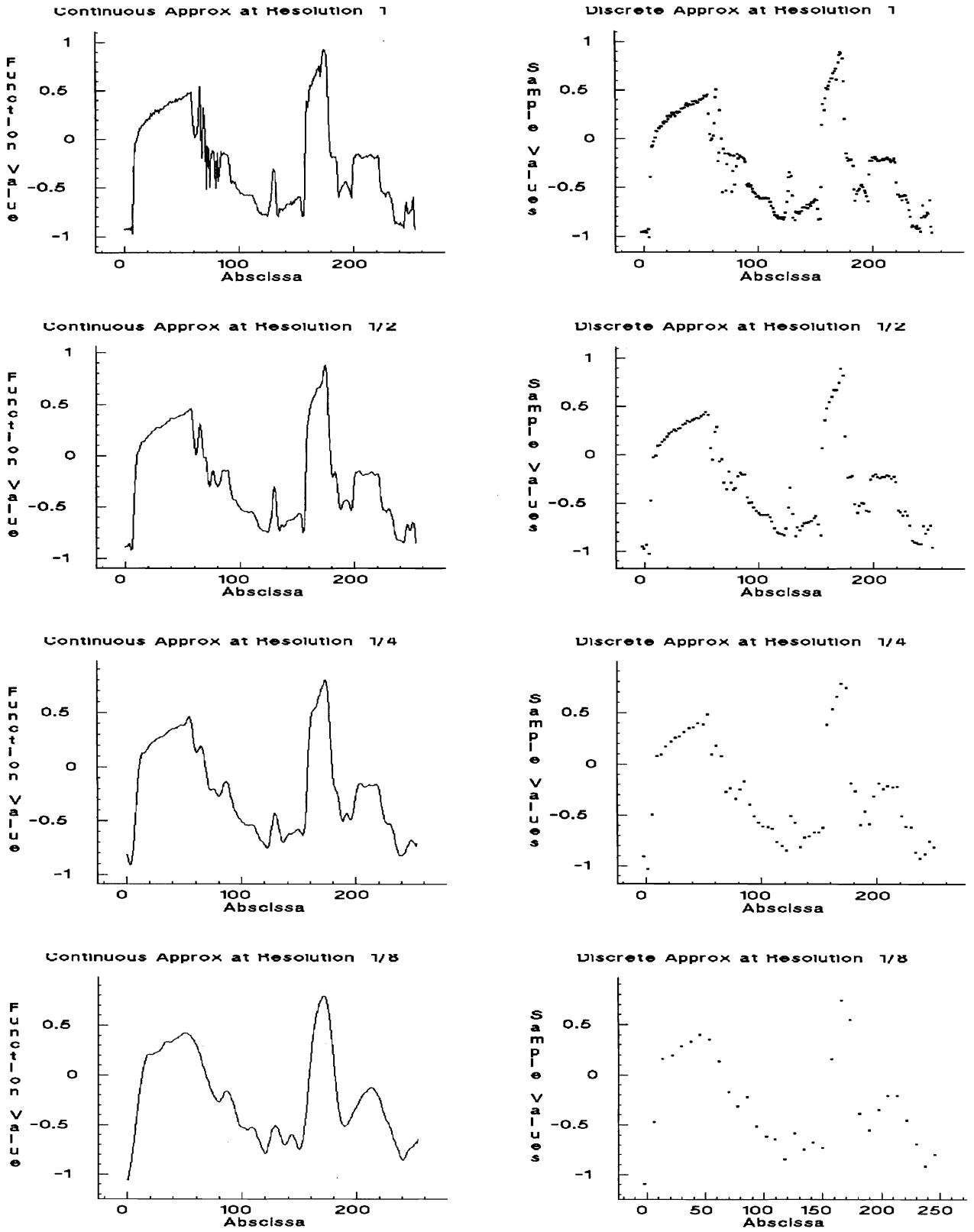


Fig. 2. *The left and right images are respectively the graphs of* S_j *and* $T_{2^j}(f)$ *for* $j = 0, -1, -2, -3$ *(resolution 1 , 1/2 , 1/4 , 1/8). When the resolution decreases smaller details gradually disappear.*

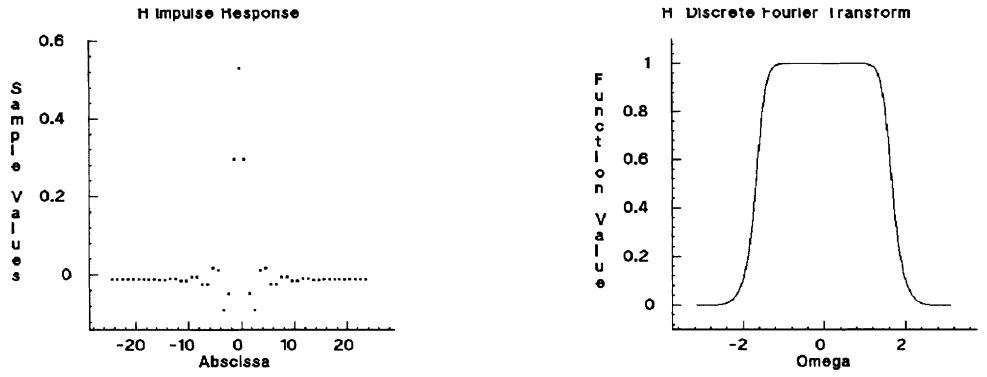


Fig. 3. *Impulse response and Discrete Fourier Transform of the filter H . This is a low pass filter and the impulse response decreases exponentially.*

Similarly to theorem 1 , theorem 3 will show that such a basis can be built by scaling and translating a unique function called a wavelet. In the following, $\bar{z}$ designates the complex conjugate of any complex number z .

Theorem 3

Let $\left(V_j\right)_{j\in \mathbf{Z}}$ be a multiresolution vector space sequence, $\phi(x)$ the scaling function and H the corresponding quadrature filter. Let $\psi(x)$ be a function whose Fourier transform is given by

$$\hat{\psi}(\omega) = e^{-i\frac{\omega}{2}} \overline{H(\tfrac{\omega}{2}+\pi)}\ \hat{\phi}(\tfrac{\omega}{2}) \tag{19}$$

Let $\psi^j(x) = \sqrt{2^j}\psi(2^j x)$: $\left(\psi^j(x-2^{-j}n)\right)_{n\in \mathbf{Z}}$ is an orthonormal basis of O_j and

$\left(\psi^j(x-2^{-j}n)\right)_{(n,j)\in \mathbf{Z}^2}$ is an orthonormal basis of $\mathbf{L}^2(\mathbf{R})$.

$\psi(x)$ is called a *wavelet.*

Appendix 4 gives some indications to prove this theorem. An orthonormal basis of O_j can thus be computed by scaling the wavelet $\psi(x)$ with a coefficient 2^j and translating it on a grid whose interval is proportional to 2^{-j} . In the frequency and the spatial domains, a wavelet decreases asymptotically like its corresponding scaling function. It can therefore have a good joint localization. The decomposition of a signal in an orthonormal wavelet basis gives an intermediate representation between a Fourier and a spatial representation. The properties of the wavelet orthonormal basis are discussed by Lemarie and Meyer [16]. From the expansion coefficients of a function in such a basis we can easily characterize its local properties such as the differentiability at any order. Fig. 4 shows the wavelet associated to the scaling function of Fig. 1, it is symmetrical with respect to ½ .

Let P_{O_j} be the orthogonal projection on the vector space O_j and $\psi_n^j(x) = \psi^j(x-2^{-j}n)$,

$$P_{O_j}(f)(x) = \sum_{n=-\infty}^{+\infty} < f\ , \psi_n^j >\ \psi_n^j(x)\ . \tag{20}$$

P_{O_j} is the detail signal at the resolution 2^{j+1} and is characterized by the set of samples :

$$D_j = \left(\sqrt{2^j} < f\ , \psi_n^j > \right)_{n\in \mathbf{Z}}.$$

We are using here the same normalizing coefficient as for S_j . D_j will be called a *discrete detail signal* at the resolution 2^{j+1} . It contains the difference of information between S_{j+1} and S_j . Similarly to (9)

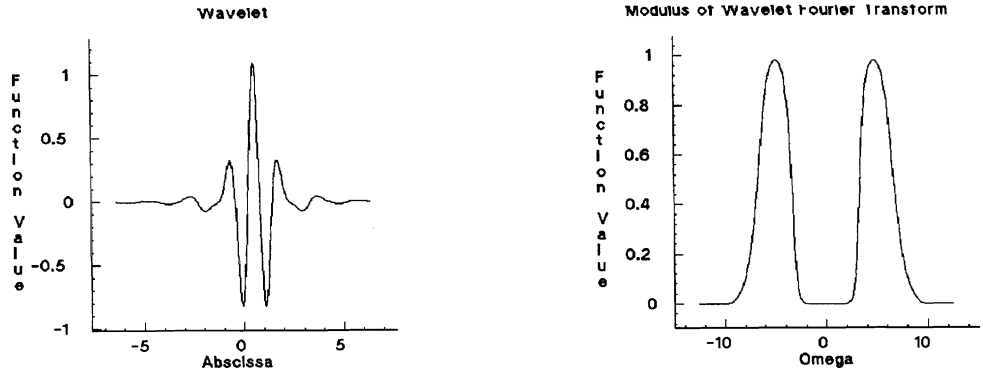


Fig. 4. *Example of wavelet with the modulus of its Fourier transform. It decreases exponentially in the spatial domain and like* $\frac{1}{\omega^4}$ *in the frequency domain.*

we can prove that

$$D_j = \left[\sqrt{2^j} \, (f(x) * \psi^j(-x))(2^{-j}n) \right]_{n \in \mathbf{Z}} . \tag{21}$$

$\psi(x)$ can also be viewed as a band pass filter and equation (19) shows that the detail signal at each resolution correspond to a particular frequency band of the signal.

We can now easily prove by induction that for any $J < 0$, the original discrete signal S_0 can be represented by the set of discrete signals

$$\left[S_J \, , (D_j)_{J \le j \le -1} \right] .$$

This is called a *wavelet representation.* It gives a reference signal at a low resolution S_J and the detail signals at the resolutions 2^j for $J \le j \le -1$. It can be interpreted as a decomposition of the original signal in an orthonormal wavelet basis or as a decomposition of the signal in a set of *independent* frequency channels like in Marr 's [18] human vision model. The independence is due to the orthogonality of the wavelet functions. However it is very difficult to have a real understanding of the model in terms of frequency decomposition because the frequency channels overlap and there is some aliasing. We can control this aliasing thanks to the orthogonality of our decomposition functions. This is why the tools of functional analysis give a better understanding of this decomposition. However, we will see that if we neglect this aliasing problem the interpretation in the frequency domain provides an intuition for understanding to the model.

3.2. Implementation of the wavelet representation

We are first going to show in this paragraph that we can compute the wavelet representation with a pyramid transform. As for S_j , D_j can indeed be easily derived from S_{j+1} .
$\psi_k^j(x) \in O_j \subset V_{j+1}$, so it can be expanded in an orthonormal basis of V_{j+1}

$$\psi_k^j(x) = \sum_{\ell=-\infty}^{+\infty} < \psi_k^j \, , \phi_\ell^{j+1} > \, \phi_\ell^{j+1}(x) \ . \tag{22}$$

Like for (11) , by changing variable in the inner product integral we can prove that :

$$< \psi_k^j \, , \phi_\ell^{j+1} > \; = \; < \psi_0^{-1} \, , \phi_{\ell-2n}^{0} > . \tag{23}$$

hence, by taking the inner product of f(x) with the functions of both sides of equation (22) we get :

$$< f \, , \psi_k^j > \; = \sum_{\ell=-\infty}^{+\infty} < \psi_0^{-1} \, , \phi_{\ell-2n}^{0} > \; < f \, , \phi_\ell^{j+1} > \ .$$

Let G be the discrete filter with impulse response

$$g(n) = \frac{1}{\sqrt{2}} < \psi_0^{-1} , \phi_n^0 > \tag{24}$$

and $\tilde{G}$ be the symmetrical filter with impulse response $\tilde{g}(n) = g(-n)$,

$$\sqrt{2^j} < f , \psi_k^j > = \sum_{k=-\infty}^{+\infty} \tilde{g}(2n-k) \sqrt{2^{j+1}} < f , \phi_k^{j+1} > . \tag{25}$$

Equation (25) shows that we can compute D_j by convolving S_{j+1} with the filter $\tilde{G}$ and keeping every other sample of the output.

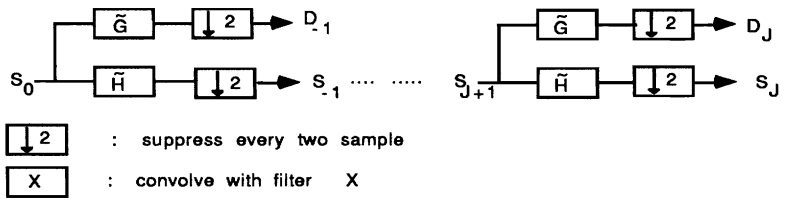


Fig. 5. *Pyramid architecture for computing the wavelet representation of a one-dimensional signal.*

Similarly to (15) , if we insert equations (23) and (24) in (22) , for j= -1 and n = 0 we have

$$\frac{1}{2} \psi(\frac{x}{2}) = \sum_{k=-\infty}^{+\infty} g(k) \phi(x-k) \tag{26}$$

Let $G(\omega)$ be the discrete Fourier transform of g(n) . The Fourier transform of equation (26) can be written :

$$\psi(2\omega) = G(\omega) \ \phi(\omega) \tag{27}$$

From equation (19) of theorem 3 we can then derive that

$$G(\omega) = e^{-i\omega} \ \overline{H(\omega+\pi)} \quad \text{and thus} \quad g(n) = (-1)^{1-n} h(1-n) \tag{28}$$

G is the *mirror* filter of H , it is a high-pass filter.

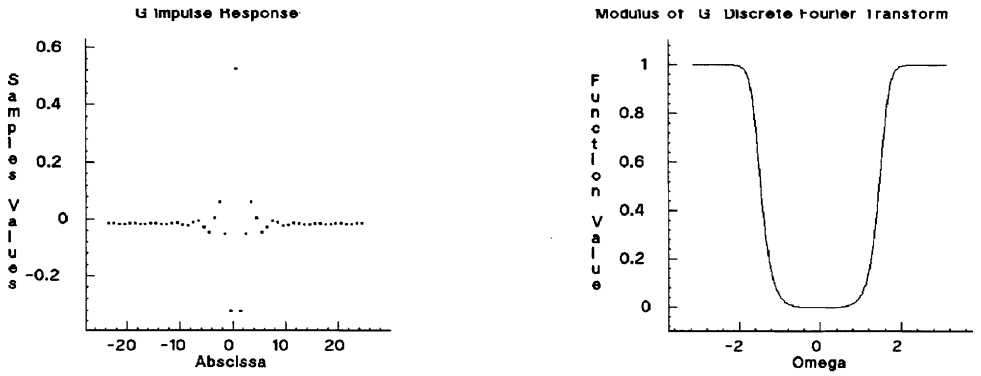


Fig. 6. *Impulse response and Discrete Fourier Transform of the filter G . This is a high pass filter and the impulse response decreases exponentially.*

Fig. 6 shows the mirror filter G of the filter H given in Fig. 3. In signal processing G and H are called [5] *quadrature mirror filters.* Equation (25) can be interpreted as a high-pass filtering of the discrete signal S_j . S_{j+1} is decomposed into S_j and D_j which respectively keeps the lower and higher frequencies. If the original signal has N samples, D_j as well as S_j will have $2^j \cdot N$ samples, so the wavelet representation

$$\left[S_J , (D_j)_{J \le j \le -1} \right]$$

has the same number of samples as S_0.

Fig. 7 shows the detail signals of the signal S_0 at the top of Fig. 2. The middle and the left graphs of Fig. 7 are respectively the discrete and continuous detail signals D_j and $P_{O_j}(f)$. We have approximated the filter G shown in Fig. 6 by taking g(n) = 0 for |n| > 8. $P_{O_j}(f)$ can be easily calculated from D_j with equation (19). On the left column of Fig. 7 we have plotted a linear interpolation of the absolute value of the detail signal samples. A signal edge contains high frequencies and thus corresponds to the amplitude peaks of the detail signals. Depending on how straight the edge is, the peak will be found at the resolution 1 or 1/2. However the highest peak at the resolution 1 is due to the "texture" between the abscissa 60 and 80 and not to an edge. This kind of problem which arises for any edge detector shows that we need a more sophisticated model to separate the edges from the rest of the signal. The differentiation between edges and textures depends on the scale of observation. What can be labeled an edge at one scale can be considered as a texture component at a smaller scale. We hence believe that a multiresolution analysis of signals can improve edge detection.

We know that the wavelet representation is complete and are now going to prove that the original discrete signal can also be reconstructed with a pyramid transform. From assertions (17) and (18) we can easily derive that $\left[\phi_k^j , \psi_k^j \right]_{n \in Z}$ is an orthonormal basis of V_{j+1}. Therefore, $\phi_k^{j+1}(x)$ can be decomposed in this basis

$$\phi_k^{j+1}(x) = \sum_{k=-\infty}^{+\infty} < \phi_k^j , \phi_n^{j+1} > \phi_k^j(x) + \sum_{k=-\infty}^{+\infty} < \psi_k^j , \phi_n^{j+1} > \psi_k^j(x) \ . \qquad (29)$$

By taking the inner product of each side of equation (29) with the function f(x) and inserting equations (10) and (23), we have

$$< f , \phi_n^{j+1} > = \sum_{k=-\infty}^{+\infty} < \phi_0^{-1} , \phi_{n-2k}^0 > < f , \phi_k^j > + \sum_{k=-\infty}^{+\infty} < \psi_0^{-1} , \phi_{n-2k}^0 > < f , \psi_k^j > \ . \qquad (30)$$

Introducing (12) and (25) in this expression, we get

$$\sqrt{2^{j+1}} < f , \phi_n^{j+1} > = 2 \left[\sum_{k=-\infty}^{+\infty} h(n-2k) \sqrt{2^j} < f , \phi_k^j > + \sum_{k=-\infty}^{+\infty} g(n-2k) \sqrt{2^j} < f , \psi_k^j > \right] \qquad (31)$$

Equation (31) shows that S_{j+1} can be reconstructed by putting zeros in between each sample of S_j and D_j, convolve them respectively with the filters H and G, add the two outputs and finally multiply the result by 2. The original discrete signal S_0 can then be reconstructed by repeating this procedure for J $\le j < 0$ as illustrated by the block diagram in Fig. 8. Fig. 9 shows the reconstruction of the discrete signal S_0 given in Fig. 2 from its wavelet representation (J = -3). By comparing the two graphs, we can see that the quality of the reconstruction is very good. The smooth as well as the most irregular parts of the signal are well rebuilt. This illustrates the numerical stability of the decomposition and reconstruction processes.

4. Extension to two dimensions

The wavelet model can be easily generalized to any dimension $n \in \mathbf{N}^*$ [11] but we will study in particular the two-dimensional case for image processing applications. The signal is now a function $f(x,y) \in \mathbf{L}^2(\mathbf{R}^2)$. We define identically a sequence of multiresolution vector spaces and the approximation of a signal f(x,y) at a resolution 2^j is still equal to its orthogonal projection on the corresponding vector space V_j. In this paper we will only study the particular case of two-dimensional scaling transformations which are built by scaling the signal both along the x and y axis. For computer implementation this will indeed allow us to compute a two-dimensional wavelet representation with *separable* filters, which reduces the computation complexity. Furthermore in such an approach we privilege the detail signal in the horizontal and vertical directions which are often the predominant directions of the image structures. Such scaling transforms are characterized by scaling functions which can be written

$$\Phi(x,y) = \phi(x)\,\phi(y)$$

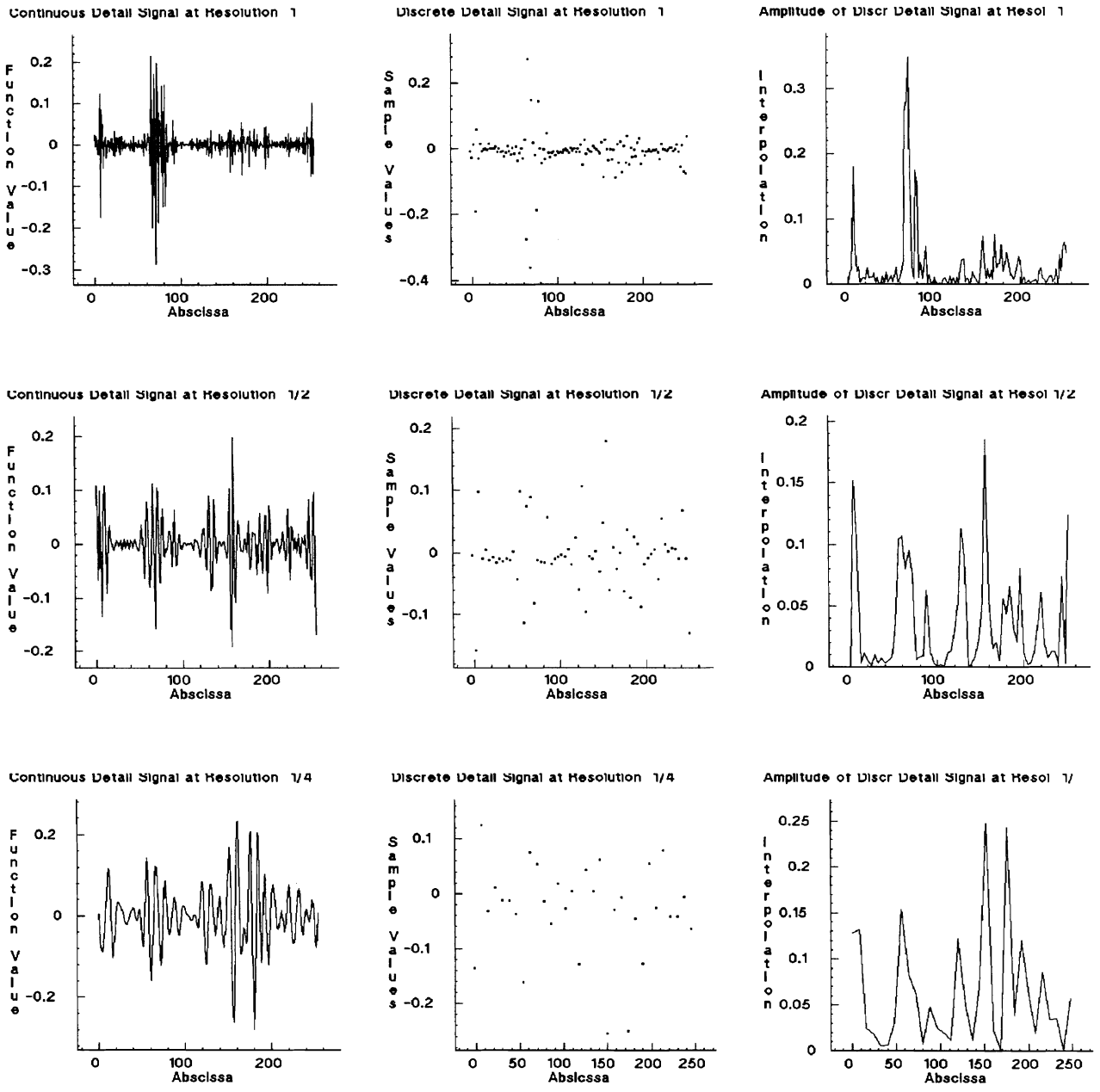


Fig. 7. *These graphs show the detail signals at the resolution 1 , 1/2 , 1/4 (j = 0 , -1 , -2). The left and middle columns are respectively the graphs of the continuous and discrete detail signals :* $P_{O_j}(f)$ *and* D_j. *The right graphs is a linear interpolation of the absolute value of the discrete detail signal samples.*

where $\phi(x)$ is a one-dimensional scaling function.

Let $\Phi^j(x,y) = 2^j\ \Phi(2^j x\ ,\ 2^j y) = \phi^j(x)\ \phi^j(y)$,

the family of functions $\left[\Phi^j(x-2^{-j}n\ ,\ y-2^{-j}m)\right]_{(n,m)\in Z^2} = \left[\phi_n^j(x)\ \phi_m^j(y)\right]_{(n,m)\in Z^2}$

is then an orthonormal basis of the vector space V_j . The approximation of a signal f(x,y) at a resolution 2^j is therefore characterized by :

$$S_j = \left[\ 2^j\ <f(x,y)\ ,\ \phi_n^j(x)\ \phi_m^j(y)>\ \right]_{(n,m)\in Z^2}.$$

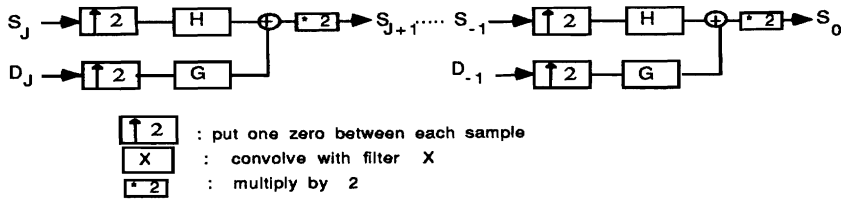


Fig. 8. *Pyramid architecture in one dimension for reconstructing the original signal from its wavelet decomposition.*

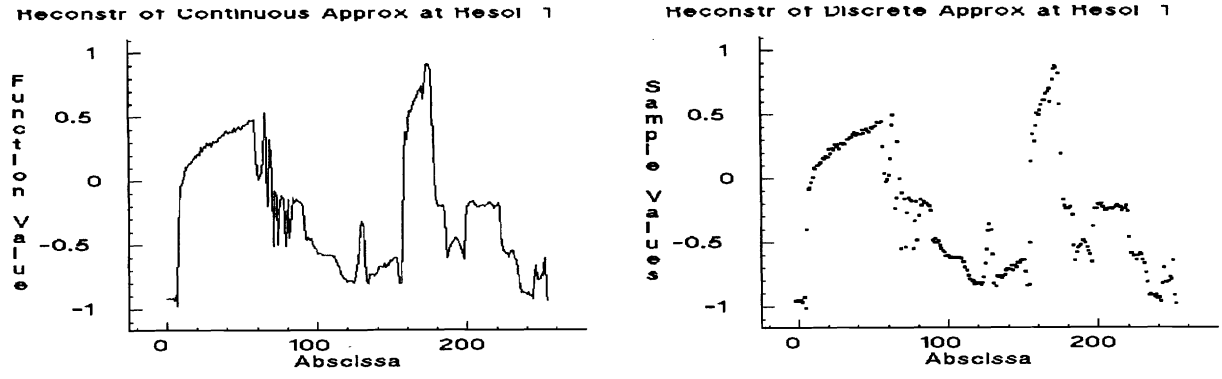


Fig. 9. *The left graph is the discrete reconstruction of the approximated signal at the resolution 1. The right graph gives the corresponding continuous signal. The quality of the reconstruction can be appreciated by comparing these graphs with the top graphs of Fig. 2.*

Similarly to (8) , in two dimensions we normalize the expansion coefficients with 2^j because

$$\int_{-\infty}^{+\infty} \Phi^j(x-2^{-j}n\ ,y-2^{-j}m)\ dx\ dy\ =\ 2^j \int_{-\infty}^{+\infty} \phi(x)\,\phi(y)\ dx\ dy\ =\ \frac{1}{2^j}\ .$$

Fig. 10 gives the discrete approximations of an image for j = 0 , -1 , -2 , -3 .

Like in the one-dimensional case, the detail signal at the resolution 2^{j+1} is equal to the orthogonal projection of the signal on a vector space O_j characterized by equations (17) and (18) . The following theorem gives a simple extension of theorem 3 and defines an orthonormal basis of O_j in order to compute the orthogonal projection.

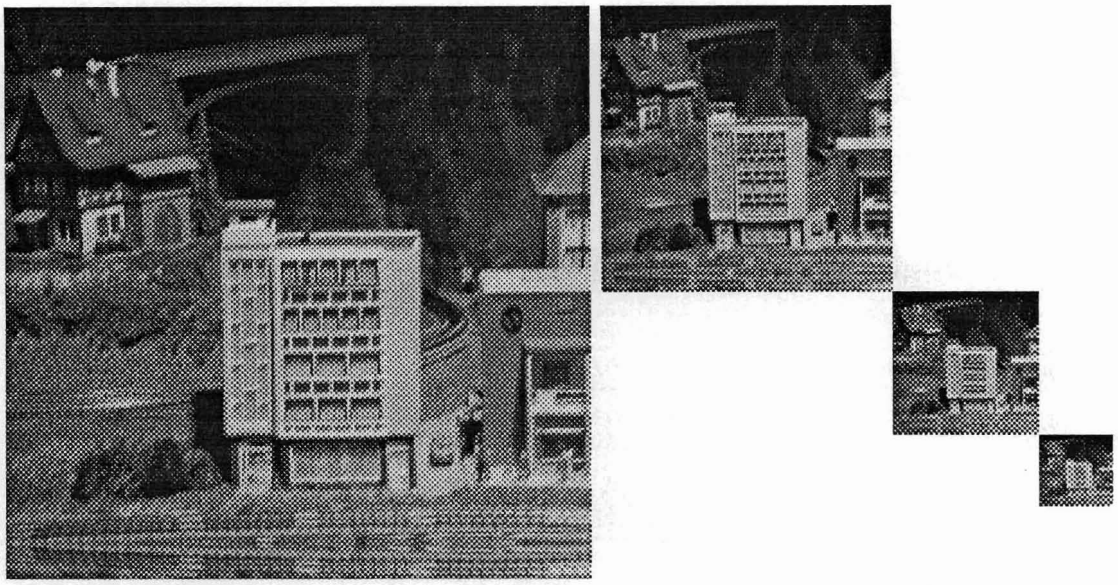


Fig. 10. *Approximations of an image at the resolution 1 , 1/2 , 1/4 , 1/8 (j = 0 , -1 , -2 , -3).*

Theorem 4

Let $\Phi(x,y) = \phi(x)\,\phi(y)$ be a two-dimensional scaling function and $\psi(x)$ the one-dimensional wavelet associated to the one dimensional scaling function $\phi(x)$. Then the family of functions

$$\left[\phi_h^j(x)\,\psi_m^j(y)\ ,\ \psi_h^j(x)\,\phi_m^j(y)\ ,\ \psi_h^j(x)\,\psi_m^j(y)\right]_{(n,m)\in Z^2}$$

is an orthonormal basis of O_j and

$$\left[\phi_h^j(x)\,\psi_m^j(y)\ ,\ \psi_h^j(x)\,\phi_m^j(y)\ ,\ \psi_h^j(x)\,\psi_m^j(y)\right]_{(n,m,j)\in Z^3}$$

is an orthonormal basis of $\mathbf{L}^2(\mathbf{R}^2)$.

Appendix 5 gives the intermediate steps of the proof. The difference of information between S_{j+1} and S_j is now given by three detail signal images:

$$D_j^1 = \left[\ 2^j < f(x,y)\ ,\ \phi_h^j(x)\psi_m^j(y) > \ \right]_{(n,m)\in Z^2}$$
$$D_j^2 = \left[\ 2^j < f(x,y)\ ,\ \psi_h^j(x)\phi_m^j(y) > \ \right]_{(n,m)\in Z^2}$$
$$D_j^3 = \left[\ 2^j < f(x,y)\ ,\ \psi_h^j(x)\psi_m^j(y) > \ \right]_{(n,m)\in Z^2}.$$

Like in the one-dimensional case, we will suppose that the output of the image digitizer is equal to S_0. For any $J < 0$, the discrete image can be thus be completely represented by the $-3J + 1$ discrete images :

$$\left[\ S_J\ ,\ (D_j^1)_{J\le j\le -1}\ ,\ (D_j^2)_{J\le j\le -1}\ ,\ (D_j^3)_{J\le j\le -1}\right].$$

This is the wavelet representation in two dimensions. S_J is the reference coarse image and the D_j^k give the detail signal for the different orientations and resolutions. If the original image has N^2 pixels the images S_{j-1} , D_{j-1}^1 , D_{j-1}^2 , D_{j-1}^3 will have each $2^j \cdot N^2$ $(j < 0)$ pixels. The total number of pixels in this new representation is equal to the number of pixels of the original image, so we do not increase the volume of data.

Similarly to (9) and (21), S_j and the signal details D_j^k can also be written

$$S_j = \left[2^j \, (f(x,y) * \phi^j(-x)\,\phi^j(-y))(2^{-j}n, 2^{-j}m) \right]_{(n,m)\in Z^2}. \tag{32}$$

$$D_j^1 = \left[2^j \, (f(x,y) * \phi^j(-x)\psi^j(-y))(2^{-j}n, 2^{-j}m) \right]_{(n,m)\in Z^2} \tag{33}$$

$$D_j^2 = \left[2^j \, (f(x,y) * \psi^j(-x)\phi^j(-y))(2^{-j}n, 2^{-j}m) \right]_{(n,m)\in Z^2} \tag{34}$$

$$D_j^3 = \left[2^j \, (f(x,y) * \psi^j(x)\psi^j(y))(2^{-j}n, 2^{-j}m) \right]_{(n,m)\in Z^2}. \tag{35}$$

The expressions (32), (33), (34) and (35) show that in two dimensions, S_j and the D_j^k are given by separable filtering of the signal along both the x and y axis. The wavelet decomposition can thus be interpreted as a signal decomposition in a set of independent, *spatially oriented* frequency channels. We can easily derive from our one-dimensional analysis that this wavelet representation can also be computed with a pyramid transform as shown in Fig. 11. It corresponds to a separable quadrature mirror filter decomposition [29]. We successively filter the rows and the columns of the S_j images at each resolution with the filters $\tilde{H}$ and $\tilde{G}$ defined previously.

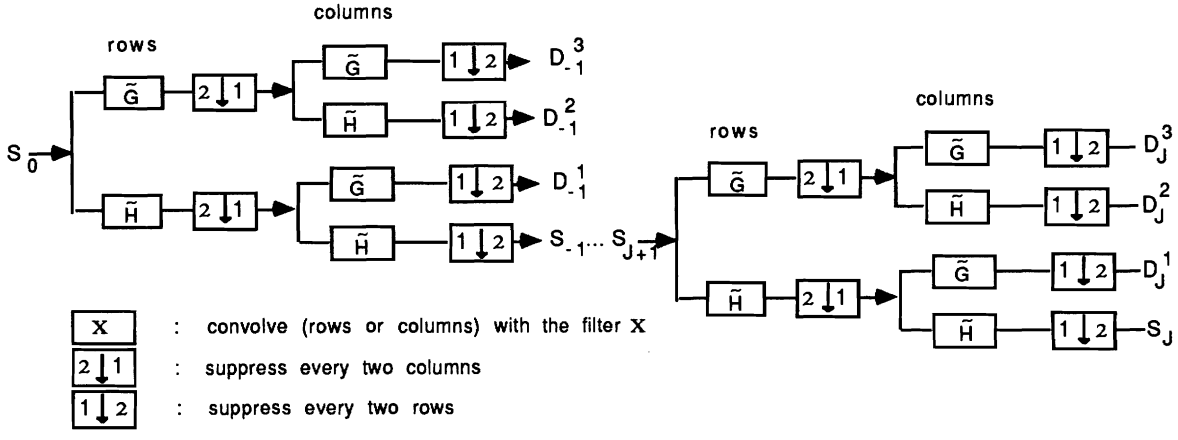


Fig. 11. *Pyramid architecture for computing the wavelet representation of a two-dimensional signal.*

Supposing that $\phi(x)$ and $\psi(x)$ are respectively a perfect low-pass and band-pass filter, Fig. 12(a) shows in the frequency domain how the discrete signal S_{j+1} is decomposed into S_j, D_j^1, D_j^2 and D_j^3. S_j corresponds to the lowest frequencies, D_j^1 gives the vertical high frequencies (horizontal edges) at the resolution 2^{j+1}, D_j^2 the horizontal high frequencies (vertical edges) and D_j^3 the high frequencies in both directions (corners). This is well illustrated by the decomposition of a white square on a black background shown in Fig. 13(b). All of the two-dimensional wavelet decompositions shown in this paper have been computed with the one-dimensional filters shown in Fig. 3 and Fig. 5. The arrangement of the D_j^k images is shown in Fig. 12(b). The black, grey and white pixels respectively correspond to negative, zero and positive coefficients. In Fig. 13(c) we have shown the absolute value of the detail signal samples. The black pixels correspond to zero whereas the white ones have the highest amplitude. As expected, the detail signal samples have high amplitude respectively on the horizontal edges, the vertical edges and the corners of the square. Fig. 14(b) shows the wavelet representation of a natural scene image for J = -3, and Fig. 14(c) gives the absolute value of the detail signal samples. The arrangement of the images is explained in Fig. 12(b). Like in one dimension, the maximum amplitude of the detail signals samples correspond to edges and rough textures area but it also provides an orientation discrimination.

The image 14(a) was provided by the Qualibrated Imaging Lab at CMU funded by the National Science Foundation.

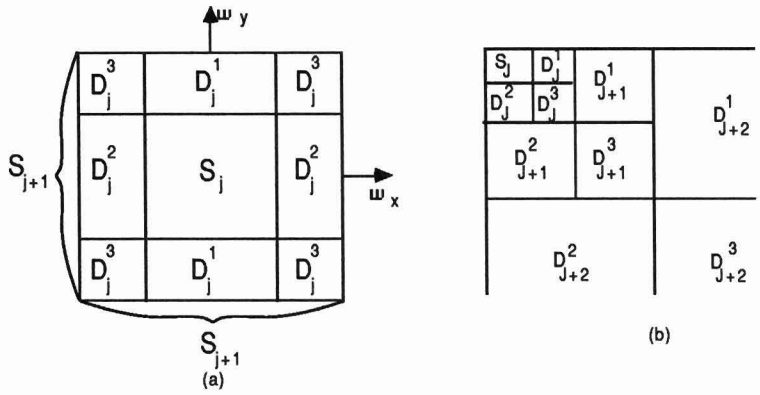


Fig. 12. *(a) Repartition in the frequency domain of the discrete images . (b) Arrangement of the* D_j^k *and* S_j *in the wavelet representation images.*

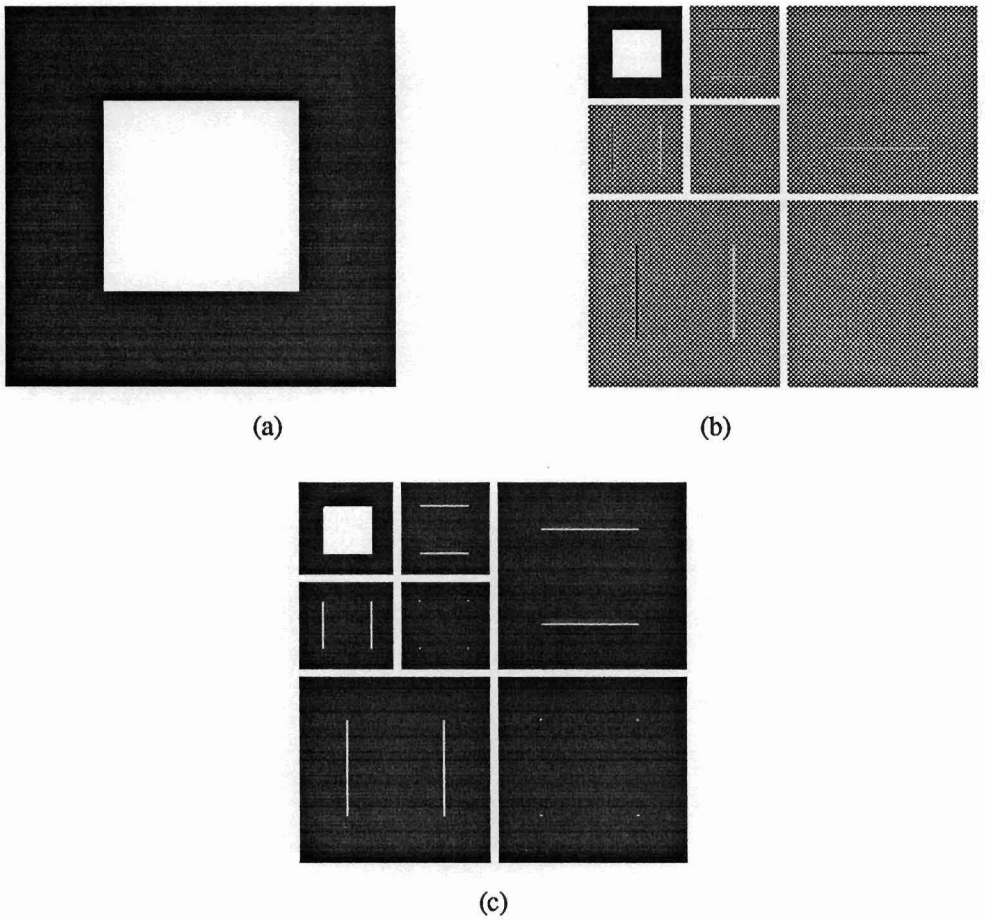


Fig. 13. *(a) Original image. (b) Wavelet representation for* $J = -2$ *. (c) Absolute value of the detail signal pixels.*

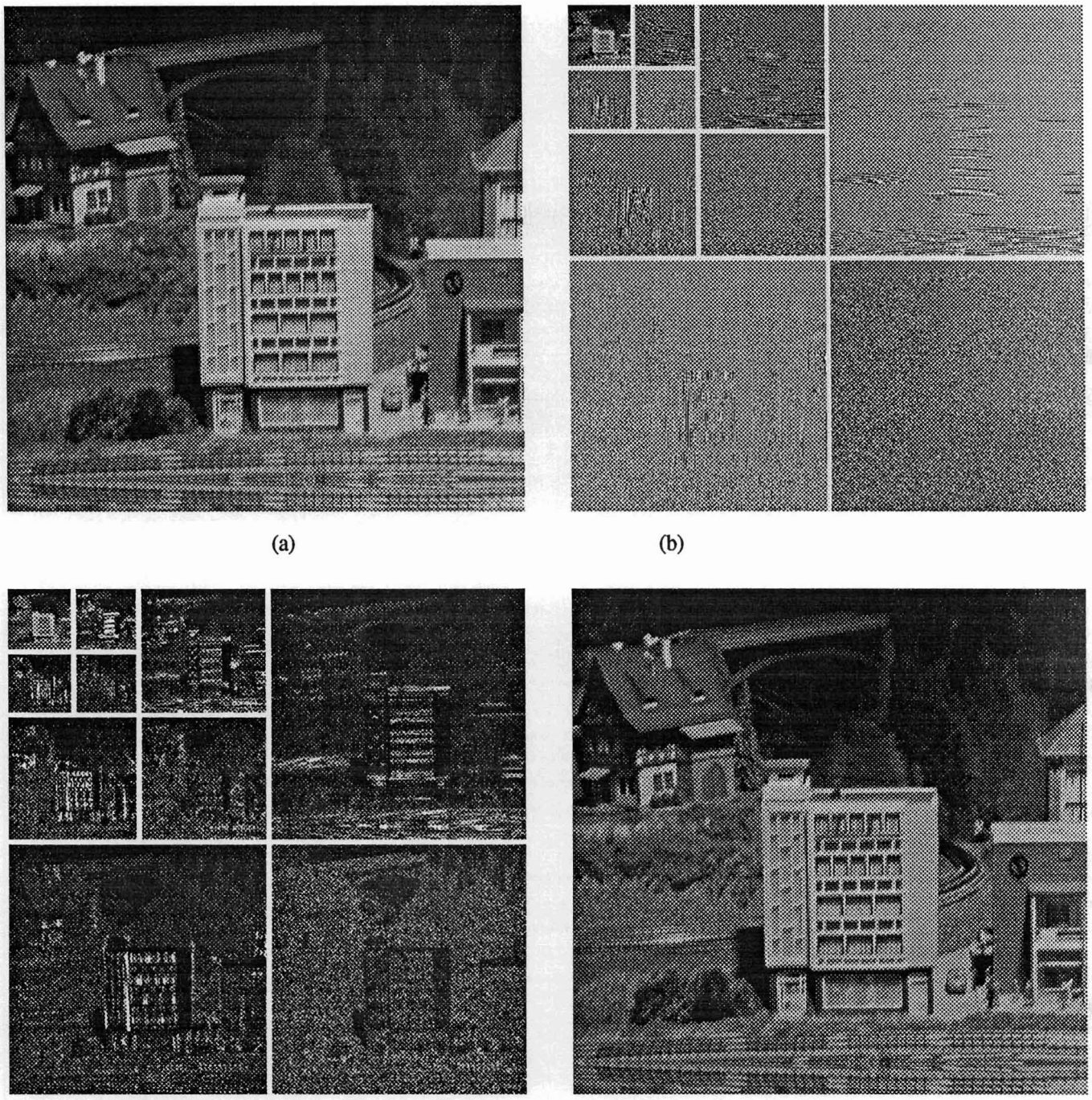


Fig. 14. *(a) Original image. (b) Wavelet representation for J = -3 . (c) Absolute value of the detail signal pixels. (d) Reconstructed image .*

We can also easily reconstruct the original signal S_0 with the pyramid architecture shown in Fig. 15. We convolve the rows and the columns of the S_j images and the detail signal images with the filters H and G defined in one dimension. Fig. 14(d) shows the image reconstructed from the wavelet representation given in Fig. 14(b). On a television monitor there are no perceptual differences between original and reconstructed images. The reconstruction process also allows us to recombine detail signal images that have undergone transformations.

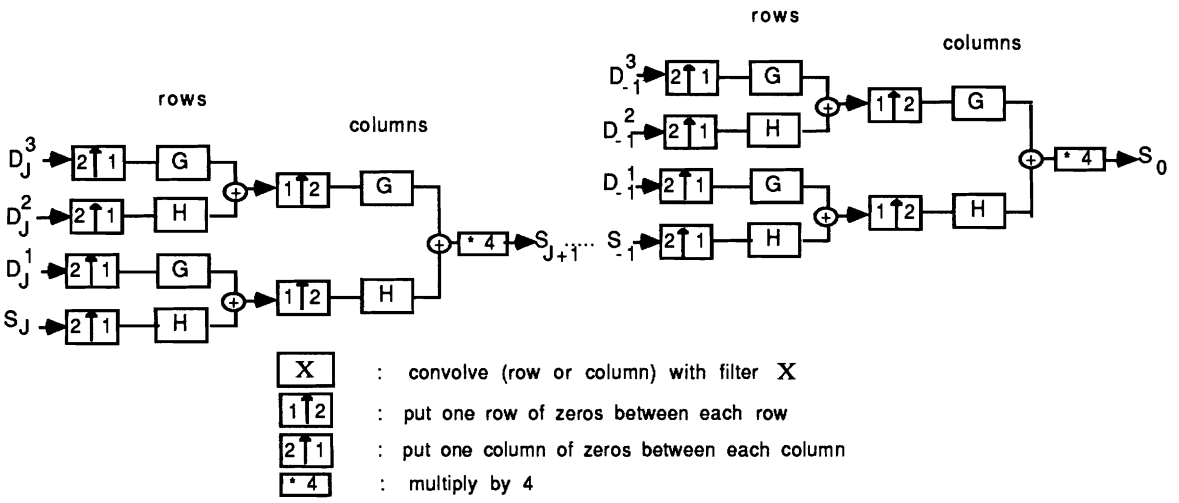


Fig. 15. *Pyramid architecture in two dimensions for reconstructing the original signal from its wavelet decomposition.*

5. Applications of the wavelet representation

In a wavelet representation the detail signals reveal some information about edges, oriented textures and gross structures which are not as easily accessible in the original image. The image is decomposed in a set of detail signals which contain the information about all the physical structures which appear specifically at a given resolution. For the separable scaling transform we have considered, the detail signal images give the horizontal, the vertical edges and the "corners" at each resolution. S_J gives locally the DC component of the image. This representation enables us to *select a priori* the relevant image information for a particular goal. Let us suppose we want to analyze a photograph taken by a satellite whose distance to the earth is known. We can compute at which resolution will appear the patterns we are looking for and therefore select the relevant detail signal images which contain the interesting information. In order to recognize a continent it is clear that we only need a very coarse description of the image which can be found in S_J and the D_j^k (k=1,2,3) for j small. Conversely, to characterize the local vegetation we just have to analyze the D_j^k for j big.

Another important application is *signal matching*. The goal is to find the position of a given pattern in an image. This kind of problem arises when we try to find some feature correspondence between several images in order to extract the depth or to measure the local motion. We can compute the wavelet representation of both the reference pattern and the image and then correlate the two representations with a coarse-to-fine strategy. First the coarser levels of the representations are put in correspondence and if the matching seems admissible we continue with the higher resolution image details. For most cases, the coarse resolution information is sufficient to eliminate the mismatches so we do not have to process the finer details, which saves computation. The orientation discrimination of our representation has some interesting applications when there is a preferred orientation in the matching problem. If we want to match a pair of stereo images for example, we know that the disparity is horizontal so the matching has to be found on a horizontal line called epipolar line. Such a matching can only be achieved by using the horizontal high frequencies (vertical edges) so we do not have to correlate the D_j^1 images.

In order to speed up the algorithms as well as save memory, it is important to reduce as much as possible the amount of data necessary to code the image. The orthogonality of the wavelet functions insures that we do not keep any redundant information, consequently we do not increase the number pixels for representing the same image. In addition, as can be seen on the images 13(b) and 14(b) , the variance of

the detail signal images is much smaller (by two or three orders of magnitude) than the variance of the original image. With standard signal processing techniques, [29], we can thus reduce considerably the number of quantization levels while having the same quantization noise. Humans have different sensitivity to the image noise depending on the frequency band where it occurs. They are most sensitive at middle frequencies and not so sensitive at very low or very high frequencies [15]. The detail signal at each resolution corresponds to a particular frequency band of the image, we can thus adapt the level of the quantization noise to the human sensitivity in the corresponding frequency band [2].

We have seen that the wavelet representation could be used for edge detection. We are now going to describe the potential applications for texture discrimination. In psychophysics Julesz [12] has shown that humans analyze textures by decomposing them with a set of basic functions called *textons*. These textons are spatially *local*, they have a particular *spatial orientation* and a narrow *frequency tuning*. The wavelet representation can also be interpreted as a texton decomposition where each texton is equal to one function of the wavelet orthonormal basis. Indeed, these functions have all the discriminative abilities required by the Julesz theory. It is easy to show that the detail signal at the resolution 2^j of a certain texture is equal to the detail signal at the resolution 2^{j+1} of the same texture seen from twice the distance. It is thus possible to analyze the texture gradient due to perspecte by comparing the detail signals when the resolution varies. In the decomposition studied in this paper we only have two orientation tunings but we could build a wavelet representation having as many orientation tunings as desired by using non-separable wavelet orthonormal bases [11]. Fig. 16(a) shows three textures synthesized by J. Beck. The human can not preattentively discriminate the middle from the right texture but can separate the left texture. In this example, the human discrimination is based mainly on the orientation of these textures as their frequency content is very similar. With a first order statistical analysis of the wavelet representation shown in Fig. 16(b), we can also discriminate the left texture but not the two others. It illustrates the ability of our representation to differentiate textures on orientation criteria. This is of course only one aspect of the problem and a more sophisticated statistical analysis is needed for modeling textures [7]. Although several psychophysical studies have shown the importance of a signal decomposition in several frequency channels [9, 1], there still is no statistical model to combine the information provided by the different channels. From this point of view, the wavelet mathematical model might be helpful to transpose some tools currently used in functional analysis to characterize the local regularity of functions.

Mandelbrot [17] has shown that some natural texture images can be modeled by fractal noise. A fractal noise F(x) is an ergodic random process which is self-similar:

$$\exists H > 0 \ , \ \forall r \in \mathbf{R} \ , \quad F(x) \text{ and } r^H F(rx) \text{ are statistically identical.}$$

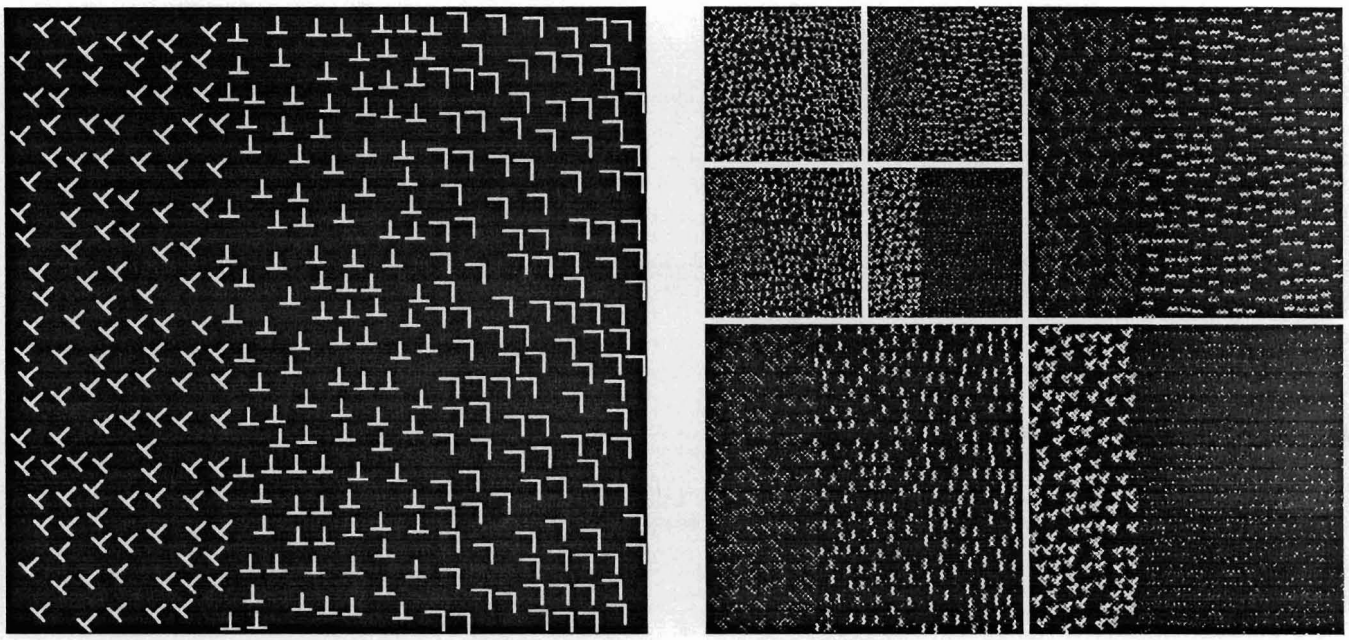
A realization of F(x) will thus look similar at any scale and for any resolution. Fractals do not provide a general model which can be extended for any kind of textures, however Pentland [24] has shown that for a fractal texture the psychophysical perception of roughness can be quantified with the fractal dimension. Fig. 17(a) shows a realization of a fractal noise which looks like a cloud. Its fractal dimension is 2.5 . Fig. 17(b) gives the absolute value of the detail signal pixels; we can see that the detail signals are similar at the different scales. The S_{-3} image gives the local DC component of the image which would correspond to the local differences of illuminations for a cloud. We are now going to show that the fractal dimension can easily be computed from the wavelet representation. We will give the proof for a one-dimensional fractal noise but it can be easily extended to two dimensions. The power spectrum of a fractal noise is given by [17]

$$P(\omega) = k \ \omega^{-2H-1}$$

The fractal dimension is related to the exponent H by

$$D = T + 1 - H \tag{36}$$

where T is the topological dimension of the space in which x varies (for images T = 2). We have seen in (21) that the detail signals D^j are obtained by filtering the signal with $\sqrt{2^j}\ \psi^j(-x)$ and sampling the output. The power spectrum of the filtered signal is



(a) (b)

Fig. 16. *(a) J.Beck textures : only the left texture is preattentively discriminable. (b) Absolute value of the wavelet representation for J = -3 : the left texture can be discriminated with a first order statistical analysis of the detail signals.*

$$P_j(\omega) = P(\omega)\ |\hat{\psi}(2^{-j}\omega)|^2 \ . \tag{37}$$

After sampling at a rate 2^j , the power spectrum is given by [22]

$$P'_j = 2^j \sum_{k=-\infty}^{+\infty} P_j(\omega + 2^j 2k\pi) \ . \tag{38}$$

Let σ_j^2 be the variance of the detail signal D_j ,

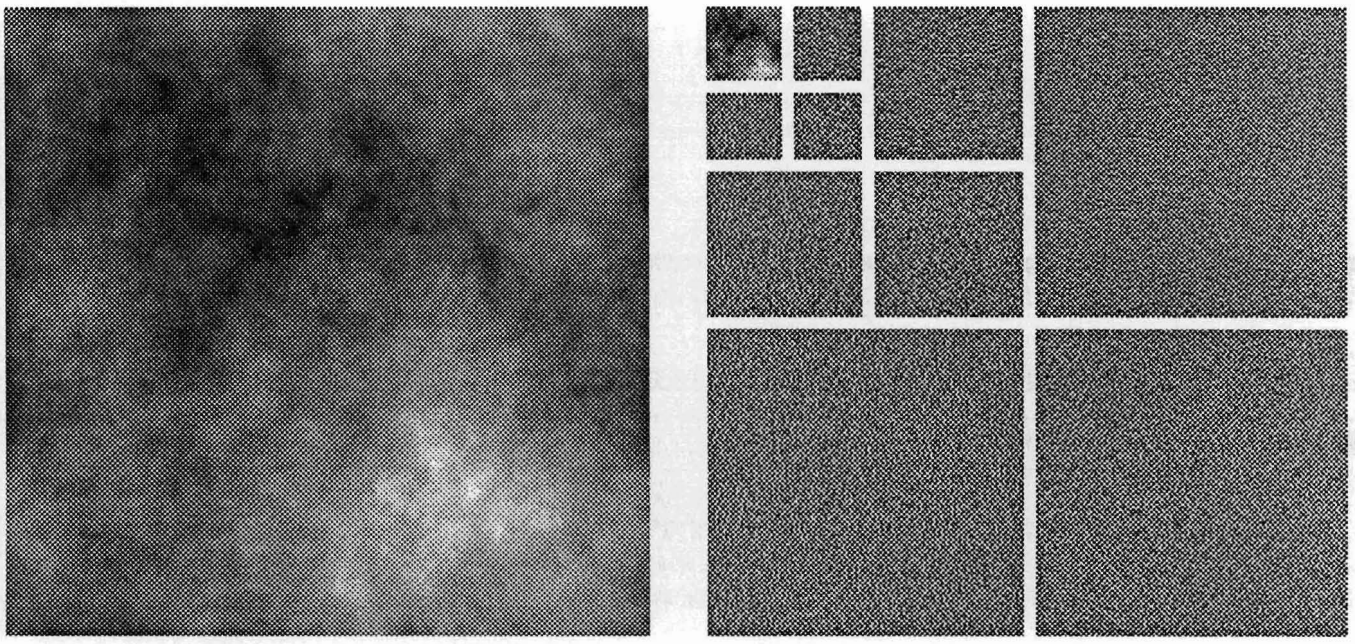
$$\sigma_j^2 = \frac{2^{-j}}{2\pi} \int_{-2^j\pi}^{+2^j\pi} P'_j(\omega)\, d\omega \ . \tag{39}$$

By inserting equations (37) and (38) in (39) and using the change of variable $\omega' = 2\omega$ in this integral we can prove that

$$\sigma_j^2 = 2^{2H}\ \sigma_{j+1}^2 \ . \tag{40}$$

For a fractal, the ratio $\frac{\sigma_j^2}{\sigma_{j+1}^2}$ should thus be constant. Equation (40) enables us to compute H and then the fractal dimension D may be computed with equation (36). This proof does not use any specific property of the wavelets and the same result can be derived with other functions such as the Gabor functions [10]. In two dimensions we can derive an equation similar to (40) for each orientation tuning of the detail signal. For the fractal shown in Fig. 17(a) we have calculated these ratios in each orientation for the first three levels. The maximum error on the fractal dimension derived from each of these ratios was 3% .

We will now briefly compare the wavelet representation with some others which have similar properties and are currently used in computer vision. The Difference Of Gaussians representation is computed by filtering the image with a set of filters equal to the difference of two gaussians of different variance. Such a representation can be interpreted as a decomposition of the signal into a set of non-oriented frequency channels and can also be built with a pyramid architecture [21]. Since the decomposition functions are not orthogonal the DOG representation increases the number of pixels by a factor of 4/3. There is no



(a) (b)

Fig. 17. *(a) Fractal noise. (b) Wavelet representation for $J = -3$. The detail signals are similar for each resolution.*

model which enables one to handle easily the redundancy of this representation, so it artificially increases computations. Moreover this representation does not have any orientation discriminability which is inconvenient for some computer vision applications such as texture segmentation.

Recently Daugmann [4] has reintroduced the Gabor representation for decomposing images. The Gabor filters are fixed variance Gaussians modulated by sinusoidal waves of different frequencies. They ate orientation selective and have an optimal joint resolution in the spatial and frequency domains and have orientation tunings. We will briefly summarize some arguments discussed by A.Grosmann [8] explaining some numerical instabilities of this representation. A Gabor decomposition is similar to a Fourier expansion but it is local because of the Gaussian window. For a signal which varies quickly with respect to the size of the window, a Gabor expansion behaves essentially like a discrete Fourier transform. It is well known that the Fourier expansion of irregular signals is very unstable because it adds many high frequency sinusoidal waves, of comparable amplitude and rapidly varying phase. The same problem appears in the Gabor expansions and J.Morlet [8] has shown for example that the Gabor coefficients could not give a robust characterization of seismic signals. In image processing such signals would correspond to the irregular textures. Some researchers [27] are now modifying the variance of the Gaussian window in the Gabor representation but it is not clear whether this new representation is complete or what its properties are.

6. Conclusion

In this paper we have described a mathematical model which enables us to understand the concept of resolution and how it relates to scale. We have seen that it is possible to compute the difference of information between different resolutions and thus define a new complete representation called the wavelet representation. It corresponds to an expansion of the original continuous signal in a wavelet orthonormal basis but can also be interpreted as a decomposition in a set of independent frequency channels with orientation tunings. The wavelet functions can be well localized both in the spatial and frequency domains so that this decomposition gives an intermediate representation between both domains. The representation is not redundant because the wavelet functions are orthogonal; it thus keeps constant the number of pixels to

code the image. The wavelet representation can be efficiently implemented with a pyramid architecture using quadrature mirror filters and the original signal can also be reconstructed with a similar architecture. The numerical stability is well illustrated by the quality of the reconstruction.

The orientation selectivity of this representation is useful for many applications. We have discussed in particular the application of the wavelet representation to signal matching, data compression, edge detection, texture discrimination and fractal analysis. Computer vision applications have been emphasized but this representation can also be helpful for pattern recognition in other domains. Grossmann and Kronland-Martinet [14] are currently working on speech recognition and J.Morlet on seismic signal analysis. The wavelet orthonormal bases are also studied both in pure and applied mathematics and have found some applications in Quantum Mechanics with the work of T. Paul [23] and P. Federbush [6].

Acknowledgments

I would to like to thank particularly R. Bajcsy for her advice and guidance throughout this research, and Y. Meyer for his help with the mathematical aspects of this paper. I am also grateful to J. Vila for his comments.

Appendices

In these appendices we will give the intermediate steps for proving the theorems without going into the mathematical details. Appendix 1 gives the proof of theorem 1 and some of the results are used in appendix 2 to prove theorem 2. Appendix 3 describes a particular class of scaling functions among which the scaling function used in this paper was chosen. Appendix 4 and 5 give the proof of theorem 3 and 4.

Appendix 1

Proof of theorem 1

We will prove theorem 1 for $j = 0$ but the result can be extended for any $j \in \mathbf{Z}$ by using equation (2). Following equation (3) and (5),

$$\exists g(x) \in V_0 \quad \text{such that} \quad I_0(g) = \varepsilon_0 \quad \text{where} \quad \varepsilon_0(k) = \begin{cases} 1 & \text{if } k = 0 \\ 0 & \text{if } k \neq 0 \end{cases}$$

Any function $f(x) \in V_0$ can thus be written

$$f(x) = \sum_{-\infty}^{+\infty} \alpha_k \, g(x-k) \tag{41}$$

Let $\hat{g}(\omega)$ and $\hat{f}(\omega)$ be the Fourier transform of g(x) and f(x),

$$\hat{f}(\omega) = M(\omega)\, \hat{g}(\omega) \quad \text{with} \quad M(\omega) = \sum_{k=-\infty}^{+\infty} \alpha_k \, e^{ik\omega} \tag{41}$$

We are looking for a function $\phi(x) \in V_0$ such that $\left[\phi(x-k)\right]_{k \in \mathbf{Z}}$ is an orthonormal family. With the Poisson formula we can show that it is equivalent to

$$\sum_{k=-\infty}^{+\infty} |\hat{\phi}(\omega + 2k\pi)|^2 = 1 \tag{42}$$

From equations (41) and (42) we can thus derive a necessary condition on $\hat{\phi}(\omega)$:

$$\hat{\phi}(\omega) = M_0(\omega)\, \hat{g}(\omega) \quad \text{with} \quad M_0(\omega) = \left[\sum_{k=-\infty}^{+\infty} |\hat{g}(\omega + 2k\pi)|^2 \right]^{-1/2} \tag{43}$$

The continuity of the isomorphism I_0 enables us to find two constants C_1 and C_2 such that

$$C_1 \leq \left[\sum_{-\infty}^{+\infty} |\hat{g}(\omega + 2k\pi)|^2 \right]^{1/2} \leq C_2 \tag{44}$$

We can then conclude from (44) that (43) defines a unique function $\phi(x)$ where $\left[\phi(x-k)\right]_{k \in \mathbf{Z}}$ is orthonormal and complete in V_0.

Appendix 2

Proof of theorem 2

(a) : $H(\omega)$ is a discrete Fourier transform and is therefore 2π periodic.

$\phi(x) = O(\frac{1}{x^3})$ so $< \phi_0^{-1}, \phi_n^0 > = O(\frac{1}{n^3})$ hence $H(\omega)$ is differentiable.

It is possible to prove that the property (6) of a multiresolution vector space sequence implies that

$$\left| \int_{-\infty}^{+\infty} \phi(x)\, dx \right| = |\hat{\phi}(0)| = 1 . \tag{45}$$

We have seen in (14) that

$$\hat{\phi}(2\omega) = H(\omega)\ \hat{\phi}(\omega) \tag{46}$$

thus $|H(0)| = 1$.

(b) : From equation (42) and (46) we can derive that

$$|H(\omega)|^2 + |H((\omega+\pi))|^2 = 1 \ .$$

Conversely we are going to show that the equation

$$\hat{\phi}(\omega) = \prod_{p=1}^{+\infty} H(2^{-p}\omega) \tag{47}$$

defines the Fourier transform of a scaling function. We need to prove that:

(α) $\hat{\phi}(\omega) \in \mathbf{L^2(R)}$ and $\left\{\phi_n^j(x)\right\}_{n\in\mathbf{Z}}$ is an orthonormal basis of a vector space V_j .

(β) $\left\{V_j\right\}_{j\in\mathbf{Z}}$ is a multiresolution vector space sequence.

Property (α) is equivalent to :

$$\forall n \in \mathbf{Z} \ , \ \int_{-\infty}^{+\infty} |\hat{\phi}(\omega)|^2 e^{ik\omega}\, d\omega = \begin{cases} 2\pi & \text{if } k = 0 \\ 0 & \text{if } k \neq 0 \end{cases} \tag{48}$$

Let us define the sequence of functions of $\mathbf{L^2(R)}$ $\left\{g_P(\omega)\right\}_{P>1}$ such that

$$g_P(\omega) = \prod_{p=1}^{P} H(2^{-p}\omega) \quad \text{for} \quad |\omega| < 2^j\pi \ , \quad \text{and} \quad \hat{g}_P(\omega) = 0 \quad \text{for} \quad |\omega| \geq 2^j\pi \ . \tag{49}$$

The sequence $\left\{g_P(\omega)\right\}_{P>1}$ converges towards $|\phi(\omega)|^2$ almost everywhere. We can easily prove that

$$\forall n \in \mathbf{Z} \ , \ \int_{-\infty}^{+\infty} g_P(\omega)\, e^{ik\omega}\, d\omega = \begin{cases} 2\pi & \text{if } k = 0 \\ 0 & \text{if } k \neq 0 \end{cases} \tag{50}$$

By using the hypothesis (c) of the theorem, it is then possible (elaborate) to apply the theorem of dominated convergence to the sequence $\left\{g_P(\omega)\right\}_{P>1}$ and therefore prove that $\phi(\omega)$ verifies (48) .

To prove that $\left\{V_j\right\}_{j\in\mathbf{Z}}$ is a multiresolution vector space sequence we need to show that the assertions (1) to (6) apply.

We can derive (46) from the hypothesis (47) . It can then be shown that $\forall (j,n) \in \mathbf{Z}^2$, $\phi_n^j(x)$ is a linear combination of the functions $\left\{\phi_n^j(x)\right\}_{n\in\mathbf{Z}}$ so $V_j \subset V_{j+1}$.

Equations (2) to (5) are not difficult to prove. (6) can be derived (elaborate) from property (45).

Appendix 3

In this appendix we describe a class of symmetrical scaling functions first discovered by P.Lemarie and G.Battle, which decrease exponentially in the spatial domain and like $\frac{1}{\omega^n}$ in the frequency domain.

Let

$$\Sigma_n(\omega) = \sum_{-\infty}^{+\infty} \frac{1}{(\omega+2k\pi)^n} \ ,$$

We can compute the close form of $\Sigma_n(\omega)$ by taking the successive derivatives of the formula

$$\Sigma_2(\omega) = \frac{1}{4\sin^2(\omega/2)} \ .$$

A Franklin scaling function of order n is defined by

$$H(\omega) = \sqrt{\frac{\Sigma_{2n}(\omega)}{2^{2n}\Sigma_{2n}(2\omega)}} \qquad \text{and therefore} \qquad \hat{\phi}(\omega) = \frac{1}{\omega^n\sqrt{\Sigma_{2n}(\omega)}} \ .$$

It is then possible to show that $\phi(x) = O(e^{-\alpha_n |x|})$ where α_n is a positive coefficient depending on n. For a Franklin scaling function of order n, the vector space V_j is the set of all functions which are n - 2 times continuously differentiable and equal to a polynomial of degree n - 1 on the intervals $[2^{-j}k\ ,\ 2^{-j}(k+1)]$. Fig. 1 shows a Franklin scaling function of order 4.

Appendix 4

Proof of theorem 3 .

This theorem will also be proved for j = 0 . We are looking for a sequence of functions $\left[\psi(x-n)\right]_{n\in\mathbf{Z}}$ which would be an orthonormal basis of O_0. We have seen in (26) that

$$\hat{\psi}(2\omega) = G(\omega)\hat{\phi}(\omega) \ . \tag{51}$$

Similarly to (42) , $\left[\psi(x-n)\right]_{n\in\mathbf{Z}}$ is orthonormal if and only if

$$\sum_{k=-\infty}^{+\infty} |\hat{\psi}(\omega+2k\pi)|^2 = 1$$

$$\text{thus} \quad |G(\omega)|^2 + |G(\omega+\pi)|^2 = 1 \ . \tag{52}$$

Because of (18) each function of the family $\left[\psi(x-n)\right]_{n\in\mathbf{Z}}$ must be orthogonal to each function of $\left[\phi(x-n)\right]_{n\in\mathbf{Z}}$. With the Poisson formula this can be written

$$\forall n \in \mathbf{Z} \quad \sum_{n=-\infty}^{+\infty} \phi(\omega+2n\pi)\ \overline{\psi(\omega+2n\pi)} = 0$$

which is equivalent to

$$H(\omega)\,\overline{G(\omega)} + H(\omega+\pi)\,\overline{G(\omega+\pi)} = 0 \ . \tag{53}$$

Conversely we can prove that the relations (51) , (52) and (53) are also sufficient for building a wavelet orthonormal basis of O_0.

A solution of (52) and (53) is given by

$$G(\omega) = e^{-i\omega}\ \overline{H(\omega+\pi)} \ .$$

From (6) , (17) and (18) one can easily derive that for $j \neq j'$, O_j is orthogonal to $O_{j'}$ and

$$\mathbf{L}^2(\mathbf{R}) = \bigoplus O_j \ .$$

Since $\left[\psi^j(x - 2^{-j}n)\right]_{n\in\mathbf{Z}}$ is an orthonormal basis of O_j , $\left[\psi^j(x - 2^{-j}n)\right]_{(n,j)\in\mathbf{Z}^2}$ is therefore an orthonormal basis of $\mathbf{L}^2(\mathbf{R})$.

Appendix 5

Proof of theorem 4

We will prove in this appendix that the family of functions :

$$\left[\phi_n^j(x)\,\psi_m^j(y)\ ,\ \psi_n^j(x)\,\phi_m^j(y)\ ,\ \psi_n^j(x)\,\psi_m^j(y)\right]_{(n,m)\in\mathbf{Z}^2} \tag{54}$$

is an orthonormal basis of O_j .

Let V_j^1 and O_j^1 be respectively the vector space generated by $\left[\phi_n^j(x)\right]_{n\in\mathbf{Z}}$ and $\left[\phi_n^j(x)\right]_{n\in\mathbf{Z}}$. Let us write $\otimes$ the tensor product of vector spaces,

$$V_{j+1} = V_{j+1}^1 \otimes V_{j+1}^1 = (O_j^1 \oplus V_j^1) \otimes (O_j^1 \oplus V_j^1)$$

hence $V_{j+1} = (V_j^1 \otimes V_j^1) \oplus (V_j^1 \otimes O_j^1) \oplus (O_j^1 \otimes V_j^1) \oplus (O_j^1 \otimes O_j^1)$.

$V_j = V_j^1 \otimes V_j^1$ thus $O_j = (V_j^1 \otimes O_j^1) \oplus (O_j^1 \otimes V_j^1) \oplus (O_j^1 \otimes O_j^1)$.

We can therefore conclude that the family (54) is an orthonormal basis of O_j.

References

[1] J. Beck, A. Sutter and R. Ivry, "Spatial frequency channels and perceptual grouping in texture segregation," *CVGIP*, vol. 3, Feb. 1987.

[2] P.J. Burt, A. E. Adelson, "The laplacian pyramid as a compact image code" *IEEE Trans. Communications* vol. 31, pp 532-540, Apr. 1983.

[3] I. Daubechies, "Orthonormal bases of compactly supported wavelets," Bell lab., 1987.

[4] J. G. Daugmann, "Six formal properties of two dimensional anisotropic visual filters. Structural principles and frequency / orientation selectivity," *IEEE Trans. System Man and Cybernetics*, vol. 13, Sept. 1983.

[5] D. Esteban and Galand C., "Applications of quadrature mirror filters to split band voice coding scheme," *Proc. ICASSP* May 1983.

[6] P. Federbush, "Quantum field theory in ninety minutes," Dept. Math. Uni. of Michigan, 1986.

[7] A. Gagalowicz, "Vers un modele de textures," These de docteur d'etat, INRIA, France, May 1983.

[8] P. Goupillaud, A. Grosmann and J. Morlet, "Cycle octave and related transform in seismic signal analysis," vol. 32, pp85-102, 1985/86.

[9] N. Graham, "Psychophysics of spatial frequency channels," pp. 215-262, Erlbaum, Hilldasle, N.J., 1981.

[10] D. Heeger and A. Pentland, "Measurement of fractal dimension using Gabor filters," Tech. Rep. TR 391, SRI AI center.

[11] S. Jaffard, P. G. Lemarie, S. Mallat and Y. Meyer, "Multiscale analysis," Centre Math., Ecole Polytechnique, Paris, 1986.

[12] B. Julesz, "Textons, the elements of texture perception and their interaction," *Nature* vol. 290, Mar. 1981.

[13] J. Koenderink "The structures of images," *Biological Cybernetics*, Springer Verlag, 1984.

[14] R. Kronland-Martinet, J. Morlet and A. Grosmann, "Analysis of sound patterns through wavelet transform," *International Journal on Pattern Analysis and Artificial Intelligence, vol. 1, Jan. 1987.*

[15] J. Kulikowski and A. Gorea, "Complete adaptation to patterned stimuli: A necessary and sufficient condition for Weber's law of contrast," *Vision Res.*, vol. 18, pp. 1223-1227, 1978.

[16] P. G. Lemarie and Y. Meyer, "Ondelettes et bases hilbertiennes," *Revista Mathematica Ibero-Americana*, vol. 2, 1986.

[17] B. Mandelbrot, *The fractal geometry of nature*, W. H. Freeman and co., New-York, 1983.

[18] D. Marr, *Vision* W. H. Freeman and co., 1982.

[19] Y. Meyer, "Principe d'incertitude, bases hilbertiennes et algebres d' operateurs," *Bourbaki seminar*, n. 662, 1985-86.

[20] P. Millar and C. Paul, "Recursive quadrature mirror filters; criteria specification and design method," *IEEE Trans. ASSP*, vol. 33, pp. 413-420, Apr. 1985.

[21] P. Burt, "Fast filter transforms for image processing," *CGIP*, vol. 16, pp. 20-51, 1981.

[22] A. Papoulis, *Probability, random variables and stochastic processes*, Mc Graw-Hill Book, New-York, 1984.

[23] T. Paul, "Affine coherent states and the radial Schrodinger equation. Radial harmonic oscillator and hydrogen atom," sent for publ.

[24] A. Pentland, "Fractal based description of natural scenes," *IEEE Trans. PAMI, vol. 6, pp. 661-674, 1986.*

[25] G. Pirani and V. Zingarelli, "Analytical formula for design of qudrature mirror filters," *IEEE Trans. ASSP, vol. 32, pp. 645-648, Jun. 1984.*

[26] M. J. Smith and T. P. Barnwell, "Exact reconstruction techniques for tree-structured subband coders," IEEE Trans. ASSP, vol. 34, Jun. 1986.

[27] M. Turner, "Texture discrimination by Gabor functions," *Biological Cybernetics*, vol. 55, pp. 71-82, 1986.

[28] A. Witkin, "Scale space filtering," *Proc. Int. Joint Conf. Articial Intell.*, 1983.

[29] J. W. Woods and S. D. O'Neil, "Subband coding of images," IEEE Trans. ASSP, vol. 34, Oct. 1986.

[30] A. Yuille and T. Poggio, "Scaling theorems for zero crossings," *IEEE Trans. PAMI*, vol. 8, Jan. 1986.