

## Wavelet array decomposition of images using a Hermite sieve

Y V VENKATESH, K RAMANI and R NANDINI

Computer Vision and Artificial Intelligence Laboratory, Department of Electrical Engineering, Indian Institute of Science, Bangalore 560012, India

**Abstract.** Generalized Hermite polynomials are used in a novel way to arrive at a multi-layered representation of images. This representation, which is centred on the creation of a new class of *wavelet arrays*, is (i) distinct from what we find in the current literature, (ii) stable, and (iii) in the manner of standard transforms, transforms the image, explicitly, into matrices of coefficients, reminiscent of Fourier series, *but at various scales*, controlled by a *scale* parameter. Among the other properties of the wavelet arrays, (a) the shape of the resolution cell in the 'phase-space' is variable even at a specified scale, depending on the nature of the signal under consideration; and (b) a systematic procedure is given for extracting the zero-crossings from the coefficients at various scales. This representation has been successfully applied to both synthetic and natural images, including textures.

**Keywords.** Image representation; wavelet transform; multi-resolution; generalized Hermite polynomials; scale space; signal decomposition; windowed Fourier transform; Fourier series; zero-crossings.

### 1. Introduction

Psycho-physical experiments have shown that biological vision employs multi-channel processing for low-level analysis of sensory data. One of the possible reasons for this type of decomposition is that the structures or details in the physical world constituting the input to the sensory system have many different sizes. In natural vision, unlike computer vision systems, the analysis, leading to information extraction and recognition of object details, seems to be independent of image scale.

Motivated by this discovery due, among others, to Hubel & Wiesel (1962), some recent investigations in the area of computer vision have dealt with the problem of representation of an image in several frequency channels. Such 'scale-space' representations have been the subject of current research.

Briefly, scale space representations and analyses are based on the idea that different characteristics of a signal reveal themselves at different levels of resolution or, equivalently, in several frequency channels. When the signal includes important structures that belong to different scales, it is useful to reorganize the signal information into a set of components of varying size. An important requirement in

any such scheme is that a small perturbation of the representation should correspond to a small modification of the original signal. At the same time, it is also desirable to localize the spatial-spectral information in the image. The need for localization of information in the spatial and frequency domains has led to signal decompositions based on either windows or frequency channels. A further requirement is that the representation should enable us to extract important properties from the image.

### 1.1 Image representation schemes

It is common practice that, given a sequence of increasing resolutions  $(r_j)_{j \in \mathbb{Z}}$ , the details of a signal at the resolution  $r_j$  are defined as the difference of information between its approximation at the resolution  $r_j$  and its approximation at the lower resolution  $r_{j-1}$ .

A structure for implementing this scheme is called the pyramid (Burt & Adelson 1983), which is a sequence of images in which each is a filtered version of its predecessor. Each image in the sequence is represented by an array which is half the size of its predecessor. The filtered signal is represented at a reduced resolution and sample densities. Assuming a 2-D function  $f(x, y)$  defined over a digital grid, one can define the pyramidal structure as a collection of sub-sampled images connected by a mapping transformation. Local operators of many scales but identical shape serve as the basis functions. The operations include low and bandpass filters and window functions. Implementation of such an approach includes blur (or reduce), expand (or make two levels of the pyramid compatible in size), and difference (or subtract) operations. However, a disadvantage of such a representation is that the elements of the image sequence are correlated.

On the hand, an approach to the extraction of localized spectral information is the use of Fourier analysis in a window of the signal. This results in a representation which is intermediate between a spatial and a frequency description. In fact, a modification was made by employing a Gaussian window in an attempt to minimize the uncertainty associated with the spatial-spectral resolution, as exemplified by the results of Marr (1982), which involve the filtering of the original image with the Laplacian of a Gaussian for various values of the variance parameter. In this case, the multiscale representation is a multichannel representation in the frequency domain where a channel corresponds to some specific bandwidth. However, the size of the resolution cell in such a representation is fixed, and, therefore, the finer details in an image when interspersed with coarse information, cannot be separated out satisfactorily.

## 2. Wavelet transforms

In order to overcome the above deficiency of the windowed Fourier transform, a combined spatial-spectral representation *à la* Gabor (1946) or the so-called wavelet transform has been proposed. As we know, the Gabor scheme uses a modulated version of the Gaussian, but, unfortunately, the Gabor functions *do not* constitute an orthogonal basis. More importantly, it is also known that they *are not* easily amenable to an orthogonalization procedure for extracting the coefficients of the signal in the Gabor space (Bastiaans 1980).

On the contrary, the wavelet transform is computed by expanding the signal into a family of functions which are the dilations and translations of a unique function,

$\phi(x)$ , called a wavelet. Grossman & Morlet (1984) decompose a function in  $L^2(\mathbb{R})$  using the family of functions  $(s^{\frac{1}{2}}\phi(sx))_{s \in \mathcal{Q}}$ . A wavelet transform is then interpreted as a decomposition of the given signal into a set of frequency channels having the same bandwidth on a logarithmic scale.

Consider a one-dimensional signal in  $L^2$ . Let  $\phi$  denote a function with sufficient decay, say  $\phi(x) \leq c/(1+x^2)$ , with

$$\int_{-\infty}^{\infty} \phi dx = 0.$$

In what follows,  $\phi_s(x)$  denotes the dilation of  $\phi(x)$  by a factor 's' and  $\phi_s^a(x)$  denotes the translation of  $\phi_s(x)$  by a factor 'a':

$$\phi_s(x) = (1/s)\phi(x/s); \quad \phi_s^a = (1/s)\phi[(x-a)/s].$$

Such a function is called a *wavelet*. The wavelet transform of a signal is then given by correlating  $f$  with  $\phi_s^a$ :

$$Wf(s, a) = \int_{-\infty}^{\infty} f(x)\phi_s^a(x)dx.$$

The choice of  $\phi$  determines the compactness of the representation, and the inversion is achieved by an appropriate inverse integral:

$$f(x) = \int_0^{\infty} \int_{-\infty}^{\infty} (1/s) Wf(s, a) \tilde{\phi}_s^a(x) da ds, \quad (1)$$

with the constraint that

$$\int_0^{\infty} (1/s) \hat{\phi}_s(u) \hat{\tilde{\phi}}_s(u) ds = 1,$$

where  $\hat{\phi}_s(u)$  is the Fourier transform of  $\phi_s(x)$ , and  $\hat{\tilde{\phi}}_s(u)$  is the Fourier transform of  $\tilde{\phi}_s(x)$ .

In practice, for computational ease,  $s$  and  $a$  are most commonly restricted to some discrete subset,  $s = 2^{-m}$  and  $a = 2^{-n}$  with  $m, n \in \mathcal{Z}$ , generating a set of dyadic wavelets. The Haar basis is a standard example for discrete wavelets.

## 2.1 Choice of the wavelet function

The choice of the wavelet function,  $\phi$ , has been the subject of many investigations. Some authors use a function which is similar to the Laplacian of the Gaussian (LoG), and others have tried to generate wavelets by recursive procedures. Common to all these attempts is the difficulty in generating orthogonal functions for a unique representation of the given signal. For instance, Mallat (1989c) starts with an orthonormal basis of  $L^2(\mathbb{R})$  generated by the family of functions

$$(\phi_2^j(x - 2^{-j}n))_{n, j \in \mathcal{Z}}.$$

However, in practice, the standard procedure for computation is a decomposition of the image using the so-called 'quadrature mirror' filters (Mallat 1989). In this approach, the signal is represented using a finite set of resolutions in powers of 2. The basic idea is to separate the higher and the lower halves of the spectrum of a signal by using second-order bandpass and low pass filters. The image is then sub-sampled corresponding to the lower half of the spectrum. This procedure is applied iteratively. This is equivalent to dividing the spectrum in successive bands, and extracting the details corresponding to these bands. Suppose, for instance, that the original signal is at resolution  $2^j$ . The result of the first band pass filtering will give us the difference of information between resolution  $2^j$  and  $2^{j-1}$ . The next band pass filtering will give us information between  $2^{j-1}$  and  $2^{j-2}$  and so on. For the sake of completeness, a detailed algorithm is given in appendix A (§ A1).

Surprisingly, the problem of *explicitly* extracting the coefficients of representation is not addressed in the literature. It may be noted that orthogonality, apart from facilitating the computation of coefficients, guarantees their uniqueness also (Daubechies 1988). In this context, other authors (see references in Mallat 1989c) have dealt with discrete wavelets  $\phi(x)$  with the property that  $((2^j)^k \phi(x))_{j \in \mathbb{Z}}$  constitute an orthonormal basis of  $L^2(\mathbb{R})$ . It is found that creating such orthonormal bases of  $L^2(\mathbb{R})$  is quite involved.

To summarize the results of the literature: (i) The procedure for the extraction of the coefficients of representation using wavelets is not given explicitly. (ii) Practical implementation using the so-called quadrature mirror filter expansion is not, strictly speaking, a wavelet coefficient representation. (iii) Most of the schemes do not lead to a unique representation, as, in fact, indicated by Daubechies (1990).

## 2.2 Features of wavelet analysis

As is well known, in Fourier analysis, for a signal that exhibits different behaviour at different spatial regions, the Fourier coefficients involve integrals defined over the entire spatial region acting as the domain of definition for the signal. Each of these coefficients reflects the presence of some frequency found somewhere in the signal. For instance, in the reconstruction of  $f(x)$ , in a region where  $f \equiv 0$ , we find that many coefficients do not vanish. Thus, it is not possible to localize the spectral content in the signal.

Wavelet analysis, like Fourier analysis, is concerned with the relation between properties of the coefficients and features in the signal. Unlike the Fourier transform, however, wavelet transform may *apparently* be used to localize the spectral components of the signal. In the discrete implementation, the support of the wavelet is chosen to be a dyadic interval. As indicated above, the Haar basis is a standard example. A window is dilated and shifted around the signal,  $f(x, y)$ . The value of the shift gives the physical location of the region under consideration, in the image. The wavelet transform of the signal, along with the wavelet coefficients in that window as support, is supposed to localize the frequency content in the signal. However, in this framework, there do exist difficulties in the choice of the size of the windows and, more importantly, in the amount of overlap permitted when the windows are shifted around.

### 3. Results based on a new wavelet array transform

Motivation for the present work comes from the fact that, in practice, signals which are spatially finite are not strictly finite in extent in the spectral domain. In view of the finite field of view of biological vision systems, and in order to develop a consistent mathematical theory, signals can be treated as though they are *approximately* infinite in both the spatial and time domains.

We employ generalized Hermite polynomials to represent the given image in multiple channels, each channel corresponding to a specific value of the scale parameter  $\sigma$ . For each channel, the representation, in contrast with the results of the literature, is *explicitly* a matrix of coefficients. Further, the number of channels is dependent on the amount of residual error permitted in the representation of the image.

Before we can state the main results, we present below in §3.1 some mathematical preliminaries, in §3.2, the proposed wavelet array, and, in §4, the properties of the new wavelet array transform.

#### 3.1 Mathematical background

Images, which are treated as 2-D functions, are assumed to be defined over  $(-\infty, \infty) \times (-\infty, \infty)$  in both the spatial and spectral domains. In what follows,  $x$  and  $y$  are independent (space) variables.

Let  $f(x, y) \in L^2(R)$  be a real-valued function of  $x, y \in \mathcal{R}$ , with the Fourier transform,

$$F(j\omega_1, j\omega_2) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) \exp(-j\omega_1 x, -j\omega_2 y) dx dy.$$

The two functions,  $f(x, y)$  and  $F(j\omega_1, j\omega_2)$ , form a Fourier integral pair. The classical uncertainty principle says that they cannot both have compact support (de Bruijn 1967, pp. 57–71; Donoho & Stark 1989).

The following concepts are needed in the determination of the size of a cell in the phase-space. The uncertainty inequality can be obtained by defining the spatial and spectral spreads of the function as follows.

The energy in a signal,  $f(x, y)$  is given by

$$E = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |f(x, y)|^2 dx dy,$$

which is the same as,

$$(1/4\pi^2) \int_{-\infty}^{\infty} |F(\omega_1, \omega_2)|^2 d\omega_1 d\omega_2.$$

Then, the effective spreads around  $(x_0, y_0)$  and  $(\omega_{10}, \omega_{20})$  are defined by

$$X_e = \left\{ [1/E] \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - x_0)^2 (y - y_0)^2 |f(x, y)|^2 dx dy \right\}^{\dagger}, \quad (2)$$

$$W_e = \left\{ [1/(4\pi^2 E)] \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (\omega_1 - \omega_{10})^2 (\omega_2 - \omega_{20})^2 |F(\omega_1, \omega_2)|^2 d\omega_1 d\omega_2 \right\}^{\frac{1}{2}}. \quad (3)$$

Using the Schwarz inequality, we obtain the standard ‘uncertainty’ inequality (de Bruijn 1967):

$$X_e W_e \geq 1/2.$$

### 3.2 Choice of basis functions and their properties

Consider the generalized 1-D Hermite polynomials parametrized by  $\sigma$  (Szegő 1975),

$$P_n(z, \sigma) = \left\{ (-1)^n \exp(x^2/2) \frac{d^n}{dx^n} [\exp(-x^2)] \right\}_{x=(z/\sigma^{1/2})}, \quad (4)$$

for  $n = 0, 1, 2, \dots$ , which can, in turn, be used to generate (by tensor product) the 2-D generalized Hermite polynomials, parametrized by  $\sigma_1$  and  $\sigma_2$ :

$$H_{m,n}(x, y, \sigma_1, \sigma_2) = P_m(x, \sigma_1) P_n(y, \sigma_2), \quad (5)$$

for  $m, n = 0, 1, 2, \dots, \infty$ .

It is known (Lebedev 1972; Higgins 1977) that the  $P_n$ 's form a complete basis for the class  $C$  of real functions,  $f(x)$ , defined on the infinite interval  $(-\infty, \infty)$ , which are piecewise continuous in every finite sub-interval  $[-a, a]$  and satisfy the condition

$$\int_{-\infty}^{\infty} (1 + x^2) \exp(-x^2/\sigma) f^2(x) dx < \infty. \quad (6)$$

For the 2-D basis representation, we use, in what follows, the vector notation,  $\mathbf{p}$ , to denote (at times, for convenience, when no explicit reference to a particular spatial variable is required) the variables  $(x, y, \sigma_1, \sigma_2)$  as a whole. The first few Hermite polynomials are:

$$\begin{aligned} H_{0,0}(\mathbf{p}) &= \exp[-(x^2/2\sigma_1) - (y^2/2\sigma_2)], \\ H_{1,0}(\mathbf{p}) &= (2x/(\sigma_1)^{\frac{1}{2}}) H_{0,0}(\mathbf{p}); \quad H_{2,0}(\mathbf{p}) = ((4x^2/\sigma_1) - 2) H_{0,0}(\mathbf{p}); \\ H_{0,1}(\mathbf{p}) &= (2y/(\sigma_2)^{\frac{1}{2}}) H_{0,0}(\mathbf{p}); \quad H_{0,2}(\mathbf{p}) = ((4y^2/\sigma_2) - 2) H_{0,0}(\mathbf{p}); \\ H_{1,1}(\mathbf{p}) &= (2x/(\sigma_1)^{\frac{1}{2}}) (2y/(\sigma_2)^{\frac{1}{2}}) H_{0,0}(\mathbf{p}); \\ H_{2,1}(\mathbf{p}) &= ((4x^2/\sigma_1) - 2) (2y/(\sigma_2)^{\frac{1}{2}}) H_{0,0}(\mathbf{p}). \end{aligned}$$

These generalized Hermite polynomials can be shown to satisfy the recurrence relations,

$$\begin{aligned} x H_{m,n}(\mathbf{p}) &= ((\sigma_1)^{\frac{1}{2}}/2) (H_{m+1,n}(\mathbf{p}) + m(\sigma_1)^{\frac{1}{2}} H_{m-1,n}(\mathbf{p})), \\ y H_{m,n}(\mathbf{p}) &= ((\sigma_2)^{\frac{1}{2}}/2) (H_{m,n+1}(\mathbf{p}) + n(\sigma_2)^{\frac{1}{2}} H_{m,n-1}(\mathbf{p})), \quad m, n = 0, 1, 2, \dots, \infty. \end{aligned}$$

These polynomials are orthogonal on  $-\infty < x, y < \infty$ . An important property of these polynomials, which facilitates multi-scale/multi-channel decomposition of

signals, is that they are Fourier transformable, and their transforms are given by

$$\tilde{H}_{m,n}(j\omega_1, j\omega_2, \sigma_1, \sigma_2) = (-j)^{m+n} H_{m,n}(\omega_1 \sigma_1, \omega_2 \sigma_2, \sigma_1, \sigma_2). \quad (7)$$

The parameters  $\sigma_1, \sigma_2$  control the effective width of the signal in both the spatial and frequency domains – the smaller the values of  $\sigma_1$  and  $\sigma_2$ , the smaller the spatial width (and greater the spectral width), and vice versa, in the directions  $x$  and  $y$  ( $\omega_1$  and  $\omega_2$ ) respectively.

In what follows,  $\Sigma_{m,n}$  denotes summation with respect to  $m$  and  $n$ , each ranging, unless otherwise indicated, from 0 to  $\infty$ . Let the  $L^2$ -norm squares of these polynomials be denoted by  $k_{m,n}$ , for  $m, n = 0, 1, 2, \dots, \infty$ .

Now we define formally the series,

$$f(x, y) = \sum_{m,n} \gamma_{m,n} H_{m,n}(\mathbf{p}), \quad -\infty < x, y < \infty, \quad (8)$$

where the coefficients  $\gamma_{m,n}$  are calculated from the relation,

$$\gamma_{m,n} = (1/k_{m,n}) \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) H_{m,n}(\mathbf{p}) dx dy, \quad m, n = 0, 1, \dots, \infty. \quad (9)$$

In practice, we use only a finite number,  $N$ , of terms. The coefficients  $\gamma_{m,n}$ , as given in (9), are obtained by the standard procedure of minimizing the mean square error. As a result, the (approximate) image reconstructed from these coefficients *will not* match with the original at all the individual points. However, theoretically, if infinite terms are used in (8), and if the real function  $f(x, y)$  defined on the infinite interval is piecewise smooth in every finite interval  $[-a, a]$ , and if the integral

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \exp[-(x^2/\sigma_1) - (y^2/\sigma_2)] f^2(x, y) dx dy$$

is finite, then the series (8) with coefficients calculated from (9) converges to  $f(x, y)$  at every continuity point of  $f(x, y)$  (Lebedev 1972; Higgins 1977).

The above expansion can also be written in matrix form:

$$f(x, y) = \mathbf{P}'(x, \sigma_1) \Gamma \mathbf{P}(y, \sigma_2), \quad -\infty < x, y < \infty, \quad (10)$$

where  $\mathbf{P}'(x, \sigma_1)$  is a row vector of 1-D Hermite polynomials,  $\Gamma$  is the coefficient matrix, and  $\mathbf{P}(y, \sigma_2)$  is the column vector of 1-D Hermite polynomials. In other words, the  $m$ th element of  $\mathbf{P}(x, \sigma_1)$  is  $P_m(x, \sigma_1)$ , the  $(m, n)$ th element of  $\Gamma$  is  $\gamma_{m,n}$ , and the  $n$ th element of  $\mathbf{P}'(y, \sigma_2)$  is  $P_n(y, \sigma_2)$ , with the prime denoting the transpose.

Assume that the given signal is expanded in terms of a finite number  $N$  of the generalized Hermite polynomials, using the scale parameters  $\sigma_{10}$  and  $\sigma_{20}$ . Then, from (10), we have

$$f_{\text{approx}}(x, y) = \mathbf{P}'(x, \sigma_{10}) \Gamma \mathbf{P}(y, \sigma_{20}), \quad -\infty < x, y < \infty, \quad (11)$$

where the dimension of the vector of polynomials  $\mathbf{P}(\cdot, \cdot)$  is  $N + 1$ .

As indicated above, in view of the fact that we have used only a finite number of terms in the representation, and the coefficients are obtained from the standard criterion of minimum mean square error, the error in the representation at scale

$(\sigma_{10}, \sigma_{20})$ , at any point  $(x, y)$ , is given by

$$\text{err}(x, y, \sigma_{10}, \sigma_{20}) = f(x, y) - f_{\text{approx}}(x, y), \quad (x, y) \in \mathcal{R}^2, \quad (12)$$

where the error is explicitly shown as dependent on  $\sigma_{10}$  and  $\sigma_{20}$ .

The Fourier transform of (12) gives

$$\text{ERR}(j\omega_1, j\omega_2, \sigma_{10}, \sigma_{20}) = F(j\omega_1, j\omega_2) - F_{\text{approx}}(j\omega_1, j\omega_2) \quad (13)$$

where

$$F_{\text{approx}}(j\omega_1, j\omega_2) = \mathbf{P}'^*(j\omega_1(\sigma_{10})^{\frac{1}{2}}, \sigma_{10}) \Gamma \mathbf{P}^*(j\omega_2(\sigma_{20})^{\frac{1}{2}}, \sigma_{20}), \\ -\infty < \omega_1, \omega_2 < \infty, \quad (14)$$

where  $\mathbf{P}^*$  indicates the Fourier transform of  $\mathbf{P}$ .

It should be noted that a minimization of the mean-square error is responsible for the inadequate representation of high frequency components in a finite expansion of the signal. The point-by-point error is an indication of the extent of the loss of frequencies higher than those of the expansion. Moreover, this error function is orthogonal to the  $N$ -dimensional vectors,  $\mathbf{P}'(x, \sigma_{10})$  and  $\mathbf{P}(y, \sigma_{20})$ . As a consequence, an expansion of  $\text{err}(x, y, \sigma_1, \sigma_2)$  using  $\sigma_1$  and  $\sigma_2$  smaller than  $\sigma_{10}$  and  $\sigma_{20}$ , respectively, can take into account the frequencies not found in  $f_{\text{approx}}(x, y)$ . Continuing in the same manner, the error in the representation at scale  $(\sigma_{11}, \sigma_{21})$ , at any point  $(x, y)$ , is then given by

$$\text{err}(x, y, \sigma_{11}, \sigma_{21}) = \text{err}(x, y, \sigma_{10}, \sigma_{20}) - \text{err}_{\text{approx}}(x, y, \sigma_{11}, \sigma_{20}), \\ (x, y) \in \mathcal{R}^2. \quad (15)$$

By combining (10)–(15), it can be shown that

$$f(x, y) = f_{\text{approx}}(x, y) + \text{err}_{\text{approx}}(x, y, \sigma_{10}, \sigma_{20}) + \text{err}_{\text{approx}}(x, y, \sigma_{11}, \sigma_{20}) + \dots \\ + \text{err}_{\text{approx}}(x, y, \sigma_{1K}, \sigma_{2L}) + \text{residual error}. \quad (16)$$

The residual error is the final error which for all practical purposes is beyond the spectral reach of the  $\sigma$ 's chosen in the multiple channels. See figure 1a for the lattice of  $\sigma$  value decomposition in the  $x$  and  $y$  directions, and figure 1b for the block diagram of the multi-layered reconstruction procedure. Appendix A (§A3) explains the notation used in figure 1b.

By virtue of the multi-stage decomposition (figure 1), the spectrum of  $f_{\text{approx}}(x, y)$  does not include that of  $\text{err}_{\text{approx}}(x, y, \sigma_{10}, \sigma_{20})$ , which in turn does not contain that of  $\text{err}_{\text{approx}}(x, y, \sigma_{11}, \sigma_{21})$ , and so on. This is equivalent to applying a sieve of Hermite polynomials – hence the title of the paper – at every level. The spectral content retained at each stage is controlled by the scale parameter,  $\sigma_{1i}, \sigma_{2j}$ .

## 4. Properties of the new wavelet array transform

### 4.1 Stability of the representation

An image  $f(x, y)$  is expanded in terms of generalized Hermite polynomials at various



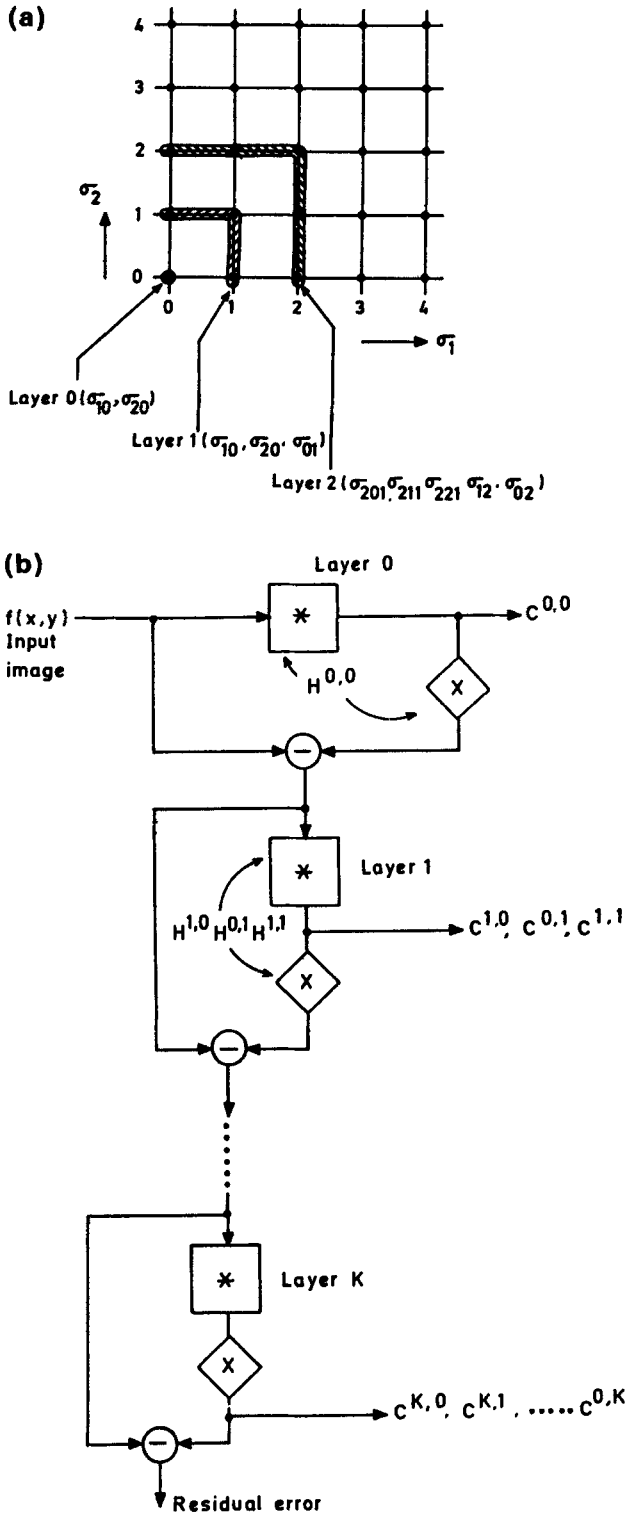


Figure 1. (a) Sampling lattice for the sigma values in the x- and y-directions. (b) Multi-layer decomposition of an image using the lattice of sigma values.

scales, denoted by  $\sigma$ , as follows:

$$f(x, y) = \sum_{m,n} \gamma_{m,n}^{00} H_{m,n}(x, y, \sigma_{10}, \sigma_{20}) + \sum_{m,n} \gamma_{m,n}^{11} H_{m,n}(x, y, \sigma_{11}, \sigma_{21}) + \dots$$

$$+ \text{residue}_{K,L}(x, y, \sigma_{10}, \sigma_{20}, \dots, \sigma_{1K}, \sigma_{2L})$$

or

$$f(x, y) = \sum_{m,n} \sum_{k,l} \gamma_{m,n}^{kl} H_{m,n}(x, y, \sigma_{1k}, \sigma_{2l}) +$$

$$+ \text{residue}_{K,L}(x, y, \sigma_{10}, \sigma_{20}, \dots, \sigma_{1K}, \sigma_{2L}), \quad (17)$$

where  $K$  and  $L$  are the number of layers of  $\sigma_1$  and  $\sigma_2$ , respectively.

The first term on the right hand side above, under the double summation symbol, is the approximation to the original signal represented using  $(K + 1)(L + 1)$  scales (layers) each having  $(N + 1)^2$  coefficients. The coefficients of the first layer are obtained from

$$\gamma_{m,n}^{00} = \langle f(x, y), H_{m,n}(x, y, \sigma_{10}, \sigma_{20}) \rangle,$$

with  $\langle \dots \rangle$  denoting inner product. Assume a perturbation,  $\delta\gamma_{m,n}^{ij}$  in one of the coefficients  $\gamma_{m,n}^{ij}$ . The resulting signal,  $f_c(x, y)$ , due to this perturbation is given by

$$f_c(x, y) = \sum_{m,n} \sum_{k,l} \gamma_{m,n}^{kl} H_{m,n}(x, y, \sigma_{1k}, \sigma_{2l}) + \sum_{m,n} \delta\gamma_{m,n}^{ij} H_{m,n}(x, y, \sigma_{1i}, \sigma_{2j}). \quad (18)$$

The change in  $f(x, y)$  due to this perturbation can be written as

$$\delta f(x, y) = f(x, y) - f_c(x, y) = \sum_{m,n} \delta\gamma_{m,n}^{ij} H_{m,n}(x, y, \sigma_{1i}, \sigma_{2j}), \quad (19)$$

from which

$$|\delta f(x, y)| \leq \max \left( |\delta\gamma_{m,n}^{ij}| \sum_{m,n} |H_{m,n}(x, y, \sigma_{1i}, \sigma_{2j})| \right)$$

or

$$|\delta f(x, y)| \leq K \sum_{m,n} |H_{m,n}(x, y, \sigma_{1i}, \sigma_{2j})|, \quad (20)$$

where  $K = \max |\delta\gamma_{m,n}^{ij}|$ .

In view of the fact that the signal  $f(x, y)$  has been assumed to be in  $L^2$  and the Hermite functions have a finite maximum for all  $m, n$  (finite),  $|\delta f(x, y)|$  is bounded. Hence, the representation is stable.

#### 4.2 Size of the phase-space resolution cell

As opposed to the wavelet transforms of the literature, the shape of the resolution cell depends on both the spatial spread and the frequency content of the image. The relations derived below determine the size of the resolution cell in 'phase-space' (Venkatesh *et al* 1991).

The effective spatial spread,  $X_{\text{spatial}}$ , in the  $x$ -direction is defined (as in (2)), by

$$X_{\text{spatial}}^2 = E(1, 0)/E(0, 0),$$

where

$$\begin{aligned} E(0,0) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f^2(x,y) dx dy \\ &= \sum_{m,n} \gamma_{m,n}^2 \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} H_{m,n}^2(x,y, \sigma_1, \sigma_2) dx dy, \end{aligned} \quad (21)$$

and  $E(1,0)$ , to be evaluated around a chosen point in the spatial domain, is given by

$$E(1,0) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^2 f^2(x,y) dx dy.$$

We employ the recurrence relation

$$xH_{m,n}(\mathbf{p}) = (\sigma_1)^{\frac{1}{2}} [mH_{m-1,n}(\mathbf{p}) + \frac{1}{2}H_{m+1,n}(\mathbf{p})],$$

which, when used in (8), gives

$$xf(x,y) \simeq \frac{1}{2}(\sigma_1)^{\frac{1}{2}} \sum_{m,n} [2(m+1)\gamma_{m+1,n} + \gamma_{m-1,n}] H_{m,n}(\mathbf{p}), \quad (22)$$

whence, we obtain

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^2 f^2(x,y) dx dy = \frac{1}{4} \sigma_1 \sum_{m,n} [2(m+1)\gamma_{m+1,n} + \gamma_{m-1,n}]^2 \times A,$$

where

$$A = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} H_{m,n}^2(x,y, \sigma_1, \sigma_2) dx dy.$$

From the above result, we obtain the effective spatial width in the x-direction as

$$X_{\text{spatial}}^2 = \left\{ \frac{1}{4} \sigma_1 \sum_{m,n} [2(m+1)\gamma_{m+1,n} + \gamma_{m-1,n}]^2 \times A \right\} / \left\{ \sum_{m,n} \gamma_{m,n}^2 \times A \right\}. \quad (23)$$

Note that (23) is to be evaluated around a chosen point in the spatial domain. On the other hand, for the computation of the effective spectral spread in the  $\omega_1$  direction, we use the recurrence relation for the Fourier transform of  $\tilde{H}_{m,n}(\mathbf{p})$ . For simplicity, we replace the argument,  $(j\omega_1, j\omega_2, \sigma_1, \sigma_2)$ , by  $(\mathbf{s})$ . We have

$$\tilde{H}_{m,n}(j\omega_1, j\omega_2, \sigma_1, \sigma_2) = (-j)^{m+n} H_{m,n}(\omega_1 \sigma_1, \omega_2 \sigma_2, \sigma_1, \sigma_2),$$

which satisfies the recursive relation

$$j\omega_1 \tilde{H}_{m,n}(\mathbf{s}) = [1/(\sigma_1)^{\frac{1}{2}}] [m\tilde{H}_{m-1,n}(\mathbf{s}) - \frac{1}{2}\tilde{H}_{m+1,n}(\mathbf{s})]. \quad (24)$$

Multiplying both sides of (8) by  $x$ , and transforming, we get

$$j\omega_1 F(j\omega_1, j\omega_2) = \sum_{m,n} \gamma_{m,n} j\omega_1 \tilde{H}_{m,n}(\mathbf{s}),$$

from which

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |j\omega_1 F(j\omega_1, j\omega_2)|^2 d\omega_1 d\omega_2 = \int_{-\infty}^{\infty} |G|^2 d\omega_1 d\omega_2, \quad (25)$$

where

$$G = \sum_{m,n} \gamma_{m,n} j\omega_1 \tilde{H}_{m,n}(s).$$

On using the recurrence relation (24), we get

$$\begin{aligned} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |j\omega_1 F(j\omega_1, j\omega_2)|^2 d\omega_1 d\omega_2 \\ = (1/4\sigma_1) \sum_{m,n} [2\gamma_{m+1,n}(m+1) - \gamma_{m-1,n}]^2 \times B, \end{aligned}$$

where

$$B = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |\tilde{H}_{m,n}(j\omega_1, j\omega_2, \sigma_1, \sigma_2)|^2 d\omega_1 d\omega_2.$$

Therefore, the effective spectral width,  $X_{\text{spectral}}$ , in the  $\omega_1$ -direction is given by

$$X_{\text{spectral}}^2 = \left[ (1/4\sigma_1) \sum_{m,n} [2(m+1)\gamma_{m+1,n} - \gamma_{m-1,n}]^2 \times B \right] / \left[ \sum_{m,n} \gamma_{m,n}^2 \times B \right]. \quad (26)$$

Note that (26) is to be evaluated around a chosen point in the spectral domain. It can be shown that

$$B = (1/\sigma_1 \sigma_2) \times A,$$

and that

$$k_{m,n}^2 = (\sigma_1 \sigma_2)^{\frac{1}{2}} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} H_{m,n}^2(x, y, 1, 1) dx dy.$$

Let

$$k'_{m,n} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} H_{m,n}^2(x, y, 1, 1) dx dy.$$

From (24) and (26), after considerable simplification, the effective space-bandwidth product (SBP), is obtained as

$$\begin{aligned} X_{\text{spatial}} \cdot X_{\text{spectral}} = \frac{1}{4} \left\{ \frac{\sum_{m,n} [2(m+1)\gamma_{m+1,n} + \gamma_{m-1,n}]^2 k'_{m,n}}{\sum_{m,n} \gamma_{m,n}^2 k'_{m,n}} \right\}^{\frac{1}{2}} \times \\ \times \left\{ \frac{\sum_{m,n} [2(m+1)\gamma_{m+1,n} - \gamma_{m-1,n}]^2 k'_{m,n}}{\sum_{m,n} \gamma_{m,n}^2 k'_{m,n}} \right\}^{\frac{1}{2}}, \quad (27) \end{aligned}$$

and the space-bandwidth ratio (SBR) as

$$\frac{X_{\text{spatial}}}{X_{\text{spectral}}} = \sigma_1 \left\{ \frac{\sum_{m,n} [2(m+1)\gamma_{m+1,n} + \gamma_{m-1,n}]^2 \times k'_{m,n}}{\sum_{m,n} [2(m+1)\gamma_{m+1,n} - \gamma_{m-1,n}]^2 \times k'_{m,n}} \right\}^{\frac{1}{2}}. \quad (28)$$

From the last result, we conclude that the SBR is directly proportional to  $\sigma_1$ , the constant of proportionality being governed by the ratio of the two quadratic forms, involving the coefficients,  $\gamma_{m,n}$ , of the expansion.

The matrices of the quadratic forms inside the square root symbol, are independent of  $\sigma_1$ . If only two coefficients are considered, it can be shown that the ratio  $\text{SBR} = \sigma_1$ . However, in general,

$$r_{\min} \sigma_1 \leq \text{SBR} \leq \sigma_1 r_{\max},$$

where  $r_{\min}$  and  $r_{\max}$  are the minimum and maximum values of the square root of the quadratic forms under the square root symbol. Similar expressions can be obtained for the effective spatial widths in the  $y$ -direction, and for the corresponding effective spectral widths in the frequency domain.

The area of the phase-space resolution cell is also dependent on the  $\gamma_{m,n}$ 's. Many possibilities arise. For instance, when  $\sigma_1$  is large (and hence the spectral window is in the low frequency part of the spectrum), and the coefficients,  $\gamma_{m,n}$ 's, are such that the term inside the square root symbol of (28) is small, then the spectral width is also small. On the contrary, when  $\sigma_1$  is small (and hence the spectral window is in the high frequency part of the spectrum), the  $\gamma_{m,n}$ 's may assume values such that the term inside the square root symbol of (28) is large. As a consequence, the spatial width could be large in the high frequency part of the spectrum.

In the classical wavelet framework, the shape of the resolution cell (which is a rectangular block with sides given by  $X_{\text{spatial}}$  and  $X_{\text{spectral}}$ ) in phase-space depends on the scale. The resolution in the spatial domain increases (decreases in the frequency domain) with an increase in the scale parameter. The area within each resolution cell is the same. In the new wavelet array framework, the shape of the resolution cell does indeed depend on the value of  $\sigma_{1i}$  and  $\sigma_{2j}$ ,  $i = 0, 1, \dots, N$  and  $j = 0, 1, \dots, M$ , but the area is independent of the location in the scale-space. See figure 2 for an illustration.

#### 4.3 Zero-crossings at various scales

The Laplacian of the Hermite polynomials at a fixed scale  $(\sigma_1, \sigma_2)$  is given by the following expression (see § A2 of appendix A),

$$\begin{aligned} [(\partial^2/\partial x^2) + (\partial^2/\partial y^2)]f(x, y) \simeq \sum_{m,n} \gamma_{m,n} \{ & (1/4\sigma_1)H_{m+2,n} + (1/4\sigma_2)H_{m,n+2} \\ & - [(1/2\sigma_1) + (1/2\sigma_2) + (m/\sigma_1) + (n/\sigma_2)]H_{m,n} \\ & + [m(m-1)/\sigma_1]H_{m-2,n} + [n(n-1)/\sigma_2]H_{m,n-2} \}. \end{aligned} \quad (29)$$

The Laplacian of the various layers of the approximated signal can be obtained merely by substituting the wavelet coefficients (at the corresponding layers). This provides an elegant way of extracting the zero-crossings at different scales of the representation.

#### 4.4 Layered decomposition

The decomposition in terms of layers of different scales has the desired property of capturing independent spectral information, by adjusting the scale parameter  $\sigma$ . By virtue of the multi-stage decomposition, the spectrum of the first-level approximation

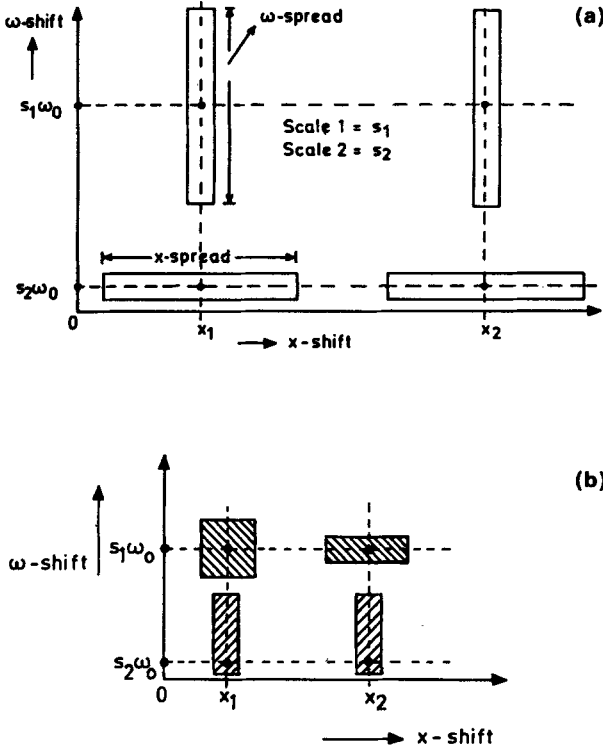


Figure 2. Resolution cell in the standard wavelet framework (a), and in the new wavelet array framework (b).

does not include that of the second level, and the second-level approximation is distinct from that of the third level, and so on.

This can be verified by using the principal component analysis of the Fourier transform of the multiple layers of reconstruction. The eigenvalues of the covariance matrix are an indication of the extent to which the individual layers contribute to the overall reconstruction. The results of computation agree with the above observations. However, we now demonstrate by a different procedure that the outputs of the various layers are indeed independent.

Let the output signals of the layers marked 0 to  $K$  and 0 to  $L$ , parametrized, respectively, by  $\sigma_{10}$  to  $\sigma_{1K}$  and  $\sigma_{20}$  to  $\sigma_{2L}$ , be given by

$$\sum_{m,n} \gamma_{m,n}^{kl} H_{m,n}(x, y, \sigma_{1k}, \sigma_{2l}), \quad k=0, 1, \dots, K \text{ and } l=0, 1, \dots, L.$$

Assume that the outputs of the layers  $(i, j)$  and  $(r, s)$ , with  $i, r=0, 1, 2, \dots, K$  and  $j, s=0, 1, 2, \dots, L$ , are *not independent*. Then for all  $(x, y)$ , we should have

$$\sum_{m,n} \gamma_{m,n}^{ij} H_{m,n}(x, \sigma_{1i}, \sigma_{2j}) = C \sum_{m,n} \gamma_{m,n}^{rs} H_{m,n}(x, y, \sigma_{1r}, \sigma_{2s}), \quad (30)$$

(where  $C$  is a constant of proportionality) from which we get

$$\sum_{m,n} \gamma_{m,n}^{ij} P_{m,n}(x, y, \sigma_{1i}, \sigma_{2j}) \exp \left[ - \left( \frac{x^2}{2\sigma_{1i}} + \frac{y^2}{2\sigma_{2j}} \right) \right] = C \left\{ \sum_{m,n} \gamma_{m,n}^{rs} P_{m,n}(x, y, \sigma_{1r}, \sigma_{2s}) \right\} \exp \left[ - \left( \frac{x^2}{2\sigma_{1r}} + \frac{y^2}{2\sigma_{2s}} \right) \right]$$

or

$$\sum_{m,n} \gamma_{m,n}^{ij} P_{m,n}(x, y, \sigma_{1i}, \sigma_{2j}) = C \left\{ \sum_{m,n} \gamma_{m,n}^{rs} P_{m,n}(x, y, \sigma_{1r}, \sigma_{2s}) \right\} \exp \left[ - \left( \frac{x^2}{2\sigma'_1} + \frac{y^2}{2\sigma'_2} \right) \right], \quad (31)$$

where  $P_n(\cdot, \cdot)$  are the polynomials associated with  $H_n(\cdot, \cdot)$ ,

$$1/\sigma'_1 = [(1/\sigma_{1i}) - (1/\sigma_{1r})]$$

and

$$1/\sigma'_2 = [(1/\sigma_{2j}) - (1/\sigma_{2s})].$$

Simplifying (31), we get

$$\sum_{m,n} \gamma_{m,n}^{ij} P_{m,n}(x, y, \sigma_{1i}, \sigma_{2j}) = C \left\{ \sum_{m,n} \gamma_{m,n}^{rs} P_{m,n}(x, y, \sigma_{1r}, \sigma_{2s}) \right\} \times \{ 1 - [(x^2/2\sigma'_1) + (y^2/2\sigma'_2)] + \frac{1}{2} [(x^2/2\sigma'_1) + (y^2/2\sigma'_2)]^2 - \dots \},$$

which after simplification, and recalling that

$$P_{0,0}(x, y, \sigma_{1i}, \sigma_{2j}) = 1; \quad P_{1,0}(x, y, \sigma_{1i}, \sigma_{2j}) = 2x/(\sigma_{1i})^{\frac{1}{2}}, \quad \text{etc.}$$

gives

$$\sum_{m,n} \beta_{mn} x^m y^n = 0, \quad (32)$$

where, for instance,

$$\beta_{00} = \sum_{m,n} \{ \gamma_{m,n}^{ij} P_{m,n}(0, 0, \sigma_{1i}, \sigma_{2j}) - K \gamma_{m,n}^{rs} P_{m,n}(0, 0, \sigma_{1r}, \sigma_{2s}) \},$$

$$\beta_{10} = \sum_{m,n} \left\{ \frac{d}{dx} [\gamma_{m,n}^{ij} P_{m,n}(x, y, \sigma_{1i}, \sigma_{2j}) - C \gamma_{m,n}^{rs} P_{m,n}(x, y, \sigma_{1r}, \sigma_{2s})] \Big|_{x=0} \right\}.$$

The polynomial (32) has only distinct contours whereas (30) is to be satisfied over the entire region. Therefore, the outputs of layers  $r$  and  $s$ , with  $r, s = 0, 1, \dots, L$ , are independent.

### 5. Comparison with the implemented results of the literature

In the proposed scheme, as implemented, each layer of the Hermite expansion is restricted to a  $13 \times 13$  matrix of coefficients, and 7 layers have been used. For simplicity, only the diagonal in the lattice of figure 1a has been computationally implemented.

The original image and the reconstructed image along with the residual error and the Fourier transform magnitude are shown in figure 3 (plates 1–3). In figure 4 (plates 3 & 4), the results of the reconstruction are shown for a texture image. Table 1 gives the coefficient matrices of the various layers for the reconstructed image.

In figure 3 (plates 1–3), it is observed that sharp boundaries (like a step edge around the eyeball in the girl image) which contain very high frequency components are not perfectly reproduced. This is to be expected as we are using a truncated expansion in terms of number of Hermite polynomials and a finite number of scales.

Further computation has been done, prompting us to the *plausible* conclusion that for a decrease in the residual error, an increase in the size of the coefficient matrix is not as effective as an increase in the number of layers.





The implementation scheme of Mallat (1989c) avoids a direct expansion in terms of coefficients, and invokes an indirect relation with the 'quadrature mirror filters'. There are many disadvantages with such an implicit wavelet scheme. One of them, for instance, is that the quadrature mirror filters cannot be used directly to obtain the zero-crossings of the decompositions. In distinct contrast, the proposed wavelet array scheme offers an explicit representation in the form of a coefficient matrix (at all levels) from which the zero-crossings can be obtained directly by synthesis.

As has been shown above, the outputs of the various layers in the proposed scheme are independent. In practice, the orthogonalization of the outputs of the layers shows that the redundancy in them is not considerable. This shows that the new scheme possesses, in addition to the multilayered structure, data compression properties also. However, a direct comparison with the discrete cosine transform (DCT), in this regard, cannot be made as the primary feature of the proposed representation is multiresolution decomposition. Besides, the DCT does not have the wavelet structure, and hence shares with the standard Fourier transform the disadvantages of 'non-localization'.

## 6. Conclusions

Representations suitable for localization of information in 2-D signals have been reviewed. A new and elegant method of signal representation based on wavelet arrays has been proposed. To this end, the use of generalized Hermite polynomials, which are orthogonal, provides a compact representation of the 2-D signal in terms of a matrix of coefficients. The novelty of the results lies in the fact that the traditional assumption of compact support in the spatial or frequency domain has been dispensed with. However, at the same time, an upper bound on the uncertainty is specified.

There do remain many problems which are unresolved.

- What is the correct or optimum sampling strategy in the  $(\sigma_1, \sigma_2)$  space and how is it related to the frequency content of the image? For instance, with a texture gradient, it has been found that the reconstruction is unsatisfactory with the sampling lattice used for other images.
- Given a physical point  $(x_p, y_p)$ , what is the frequency content of the image in its neighbourhood?
- Conversely, where do we find, in the physical image, regions with a prescribed frequency content?

The authors thank the reviewer(s) for some helpful suggestions which led to the present improved version of the paper. The authors are grateful to Professor T Kailath, Stanford University, Stanford, California, for his support, encouragement and keen interest in the work.

## Appendix A

### A1. Implementation of the quadrature-mirror filter

The quadrature-mirror can be implemented (Mallat 1989c) using a pyramidal structure. A brief algorithm is given below.

Let  $S^j$  be the original signal at resolution  $2^j$ . Let  $U_{\Pi/2}$  be the low pass filter with a cutoff frequency  $\Pi/2$  and  $G_{\Pi/2}$  be the high pass filter with a cutoff frequency of  $\Pi/2$ .

*Decomposition:*

- Filter  $S^j$  by  $G_{\Pi/2}$  to get  $W^j$ .
- Expand  $W^j$  by a factor of 4 by interpolating between samples.
- Filter  $S^j$  by  $U_{\Pi/2}$  to get  $S^{j-1}$ .
- Filter  $S^{j-1}$  by  $G_{\Pi/2}$  to get  $W^{j-1}$ .
- Expand  $W^{j-1}$  by a factor of 2 by interpolating between samples.
- Filter  $S^{j-1}$  by  $U_{\Pi/2}$  to get  $S^{j-2}$ .
- Filter  $S^{j-2}$  by  $G_{\Pi/8}$  to get  $W^{j-2}$ .
- Filter  $S^{j-2}$  by  $U_{\Pi/8}$  to get  $S^{j-3}$ . Sample  $S^{j-3}$  by 2 by omitting one point between samples.
- Filter  $S^{j-2}$  by  $G_{\Pi/8}$  to get  $W^{j-2}$ .
- Filter  $S^{j-2}$  by  $U_{\Pi/8}$  to get  $S^{j-3}$ . Sample  $S^{j-3}$  by 2 by omitting one point between samples.

This completes 4 levels of resolution.

*Reconstruction:*

- Expand  $S^{j-4}$  by a factor of 2 to get  $S_{\text{exp}}^{j-4}$ .
- Filter  $S_{\text{exp}}^{j-4}$  by  $U_{\Pi/8}$  and filter  $W^{j-3}$  by  $G_{\Pi/8}$ . Add up the results to get  $S^{j-3}$ .
- Expand  $S^{j-3}$  by a factor of 2 to get  $S_{\text{exp}}^{j-3}$ .
- Filter  $S_{\text{exp}}^{j-3}$  by  $U_{\Pi/8}$  and filter  $W^{j-2}$  by  $G_{\Pi/8}$ . Add up the results to get  $S^{j-2}$ .
- Reduce  $W^{j-1}$  by a factor of 2 to get  $W_{\text{red}}^{j-1}$ .
- Filter  $S^{j-2}$  by  $U_{\Pi/4}$  and filter  $W_{\text{red}}^{j-1}$  by  $G_{\Pi/4}$ . Add up the results to get  $S^{j-1}$ .
- Reduce  $W^j$  by a factor of 4 to get  $W_{\text{red}}^j$ .
- Filter  $S^{j-1}$  by  $U_{\Pi/2}$  and filter  $W_{\text{red}}^j$  by  $G_{\Pi/2}$ . Add up the results to get  $S^j$ .

In the above algorithm for reconstruction, there is *no* reference to the extraction of the wavelet coefficients. Further, image properties/features (like the zero-crossings) of the original image cannot be directly extracted from the filter outputs.

## A2. Zerocrossings at various scales

From (8), we get

$$\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right)f(x, y) \simeq \sum_{m,n} \gamma_{m,n} \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right)H_{m,n}. \quad (\text{A1})$$

In what follows, we omit, for simplicity, the arguments  $(x, y, \sigma_1, \sigma_2)$  in the  $H_{m,n}$ 's.

From the recursive relation,

$$\frac{\partial H_{m,n}}{\partial x} = \frac{x}{\sigma_1} H_{m,n} - \frac{1}{(\sigma_1)^{\frac{1}{2}}} H_{m+1,n},$$

we derive

$$\frac{\partial^2}{\partial x^2} H_{m,n} = \left(\frac{1}{\sigma_1} + \frac{x^2}{\sigma_1^2}\right) H_{m,n} - \frac{2x}{\sigma_1(\sigma_1)^{\frac{1}{2}}} H_{m+1,n} + \frac{1}{\sigma_1} H_{m+2,n};$$

and

$$\frac{\partial^2}{\partial y^2} H_{m,n} = \left( \frac{1}{\sigma_2} + \frac{y^2}{\sigma_2^2} \right) H_{m,n} - \frac{2y}{\sigma_2(\sigma_2)^{\frac{1}{2}}} H_{m,n+1} + \frac{1}{\sigma_2} H_{m,n+1}.$$

Therefore,

$$\begin{aligned} \left( \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \right) H_{m,n} &= \left( \frac{1}{\sigma_1} + \frac{1}{\sigma_2} + \frac{x^2}{\sigma_1^2} + \frac{y^2}{\sigma_2^2} \right) H_{m,n} - \frac{2x}{\sigma_1(\sigma_1)^{\frac{1}{2}}} H_{m+1,n} \\ &\quad - \frac{2y}{\sigma_2(\sigma_2)^{\frac{1}{2}}} H_{m,n+1} + \frac{1}{\sigma_1} H_{m+2,n} + \frac{1}{\sigma_2} H_{m,n+2}. \end{aligned} \quad (\text{A2})$$

Further, from the recursive relation,

$$xH_{m,n} = (\sigma_1)^{\frac{1}{2}} \left[ mH_{m-1,n} + \frac{1}{2}H_{m+1,n} \right],$$

we have

$$x^2 H_{m,n} = (\sigma_1/4)H_{m+2,n} + (\sigma_1/2)(2m+1)H_{m,n} + m(m-1)\sigma_1 H_{m-2,n}.$$

Similarly,

$$y^2 H_{m,n} = (\sigma_2/4)H_{m,n+2} + (\sigma_2/2)(2n+1)H_{m,n} + n(n-1)\sigma_2 H_{m,n-2}.$$

Therefore,

$$\begin{aligned} \left( \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \right) H_{m,n} &= \frac{1}{4\sigma_1} H_{m+2,n} + \frac{1}{4\sigma_2} H_{m,n+2} - \left( \frac{1}{2\sigma_1} + \frac{1}{2\sigma_2} + \frac{m}{\sigma_1} + \frac{n}{\sigma_2} \right) H_{m,n} \\ &\quad + \frac{m(n-1)}{\sigma_1} H_{m-2,n} + \frac{n(n-1)}{\sigma_2} H_{m,n-2}. \end{aligned} \quad (\text{A3})$$

Using this in (30), we get for any layer,

$$\begin{aligned} \left( \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \right) f(x,y) &\simeq \sum_{m,n} \gamma_{m,n} \left[ \frac{1}{4\sigma_1} H_{m+2,n} + \frac{1}{4\sigma_2} H_{m,n+2} \right. \\ &\quad \left. - \left( \frac{1}{2\sigma_1} + \frac{1}{2\sigma_2} + \frac{m}{\sigma_1} + \frac{n}{\sigma_2} \right) H_{m,n} + \frac{m(m-1)}{\sigma_1} H_{m-2,n} + \frac{n(n-1)}{\sigma_2} H_{m,n-2} \right] \\ &\equiv \sum_{m,n} \mu_{m,n} H_{m,n}, \end{aligned} \quad (\text{A4})$$

where

$$\begin{aligned} \mu_{m,n} &= \sum_{m,n} \frac{1}{4\sigma_1} \gamma_{m-2,n} + \frac{1}{4\sigma_2} \gamma_{m,n-2} - \left( \frac{1}{2\sigma_1} + \frac{1}{2\sigma_2} + \frac{m}{\sigma_1} + \frac{n}{\sigma_2} \right) \gamma_{m,n} \\ &\quad + \frac{(m+2)(m+1)}{\sigma_1} \gamma_{m+2,n} + \frac{(n+2)(n+1)}{\sigma_2} \gamma_{m,n+2}. \end{aligned}$$

From (32) the Laplacian of the approximated signal can be obtained in the various layers, merely by substituting the wavelet coefficients for the corresponding layers.

## A3. Notation for figure 1

$H^{k,l}$  is an  $(N + 1) \times (N + 1)$  matrix.

$$\{H_{m,n}(x, y, \sigma_{1k}, \sigma_{2l})\},$$

$m, n = 0, 1, 2, \dots, N; k = 0, 1, 2, \dots, K; l = 0, 1, 2, \dots, L.$

'\*' denotes inner product for coefficient extraction and 'x' denotes matrix multiplication of the type (11) for reconstruction.  $C^{k,l}$  is an  $(N + 1) \times (N + 1)$  matrix of coefficients at the  $(k, l)$ th layer and  $\text{err}^{k,l}(x, y)$  is the error at the end of the  $(k, l)$ th layer as a summation of outputs of the previous layers.

## References

- Bastiaans M J 1980 Gabor's signal expansion and degrees of freedom of a signal. *Proc. IEEE* 68: 538–539
- Burt P J, Adelson E H 1983 The Laplacian pyramid as a compact image code. *IEEE Trans. Commun.* COM-31: 532–540
- Daubechies I 1988 Orthonormal basis of compactly supported wavelets. *Commun. Pure Appl. Math.* 41: 909–996
- Daubechies I 1990 The wavelet transform time-frequency localization and signal analysis. *IEEE Trans. Inf. Theory* IT-36: 961–1004
- de Bruijn N G 1967 *Uncertainty principles in Fourier analysis in inequalities* (ed.) O Shisha (New York: Academic Press)
- Donoho D L, Stark P B 1989 Uncertainty principles and signal recovery. *SIAM J. Appl. Math.* 49: 906–931
- Gabor D 1946 Theory of communication. *J. Inst. Elec. Eng. (London)*, 93: 429–457
- Grossman A, Morlet J 1984 Decomposition of Hardy functions into square integrable wavelets of constant shape. *SIAM J. Math.* 15: 723–736
- Higgins J R 1977 *Completeness and basis properties of sets of special functions.* (Cambridge: University Press)
- Hubel D H, Wiesel T N 1962 Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *J. Physiol.* 160: 106–154
- Lebedev N N 1972 *Special functions and their applications* (New York: Dover) pp. 60–67
- Mallat S 1989a Multi-resolution approximation and wavelet orthonormal based of  $L^2$ . *Trans. Am. Math. Soc.* 3(15): 69–87
- Mallat S 1989b A theory for multiresolution signal decomposition – the wavelet representation. *IEEE Trans. Pattern Anal. Mach. Intell.* PAMI-11: 674–693
- Mallat S 1989c Multi-frequency channel decompositions of images and wavelet models. *IEEE Trans. Acoust. Speech Signal Process.* ASSP-37: 2091–2110
- Marr D 1982 *Vision* (San Francisco, CA: Freeman)
- Szego G 1975 *Orthogonal polynomials* (Providence, NJ: Am. Math. Soc.)
- Venkatesh Y V, Nandini R, Ramani K 1991 Vector wavelet decomposition of 1-D signals, Technical Report, Department of Elec. Engg. IISc, July 1991 and IAPR International conference paper (Aug–Sep 1992)

Figure 3. (a-e)

(a)

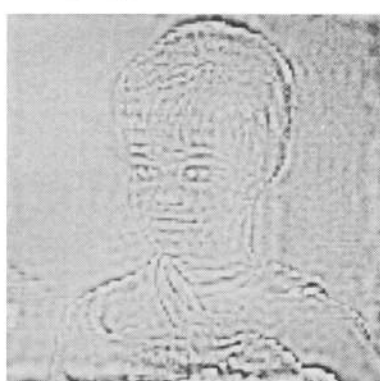


Plate 1

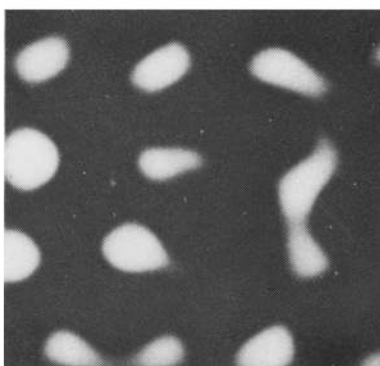
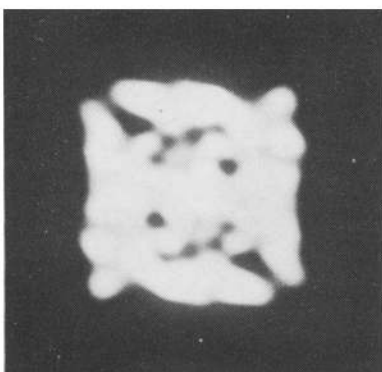
(b)



(c)



(d)



(e)

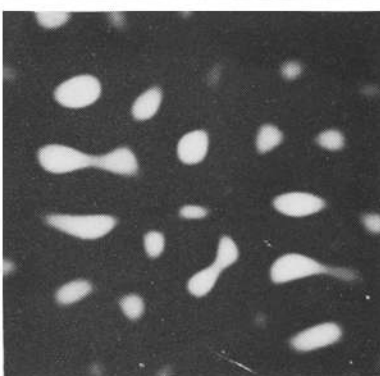
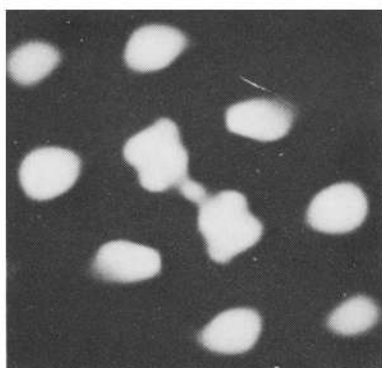


Figure 3. (f-i)

Plate 2

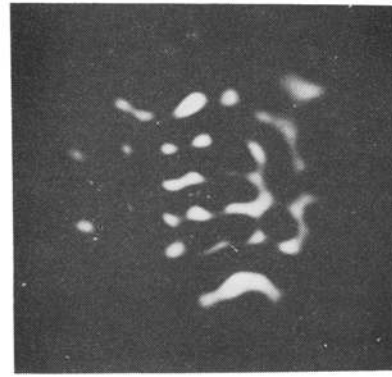
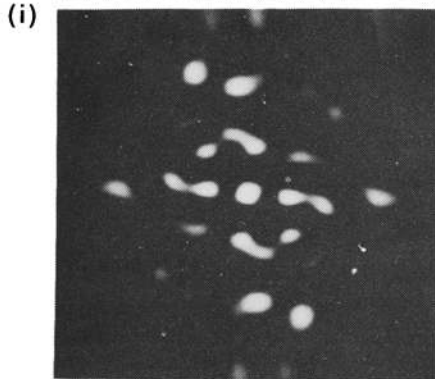
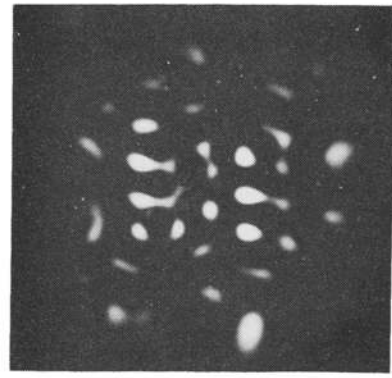
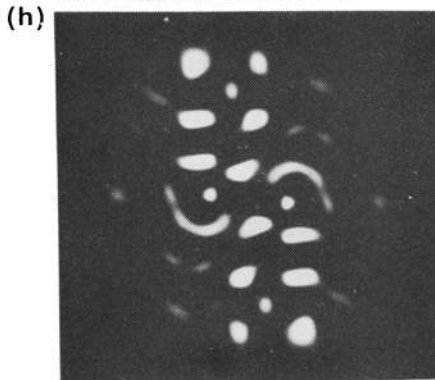
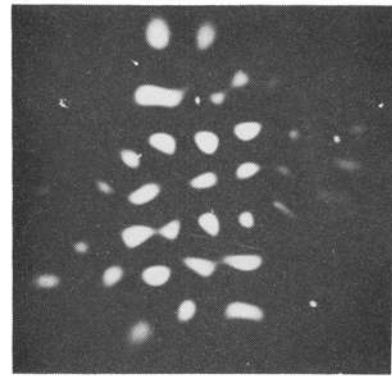
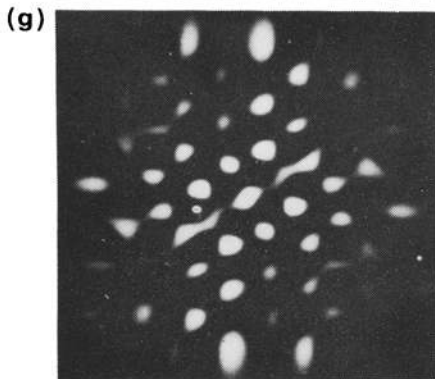
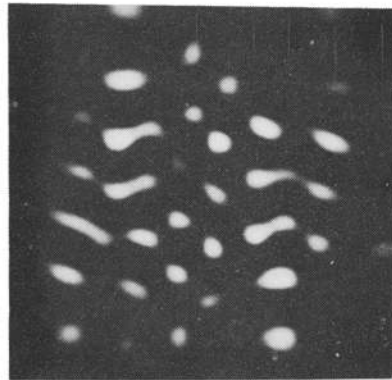
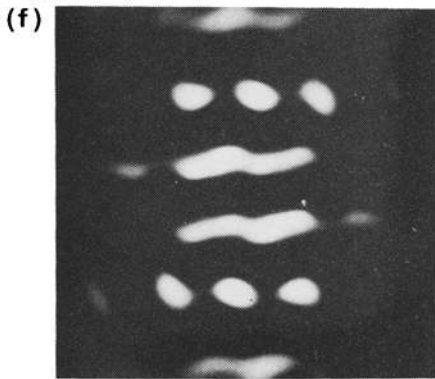
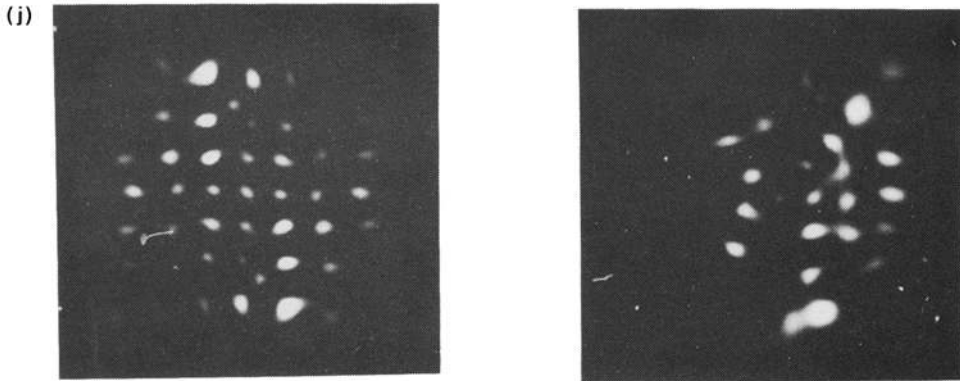
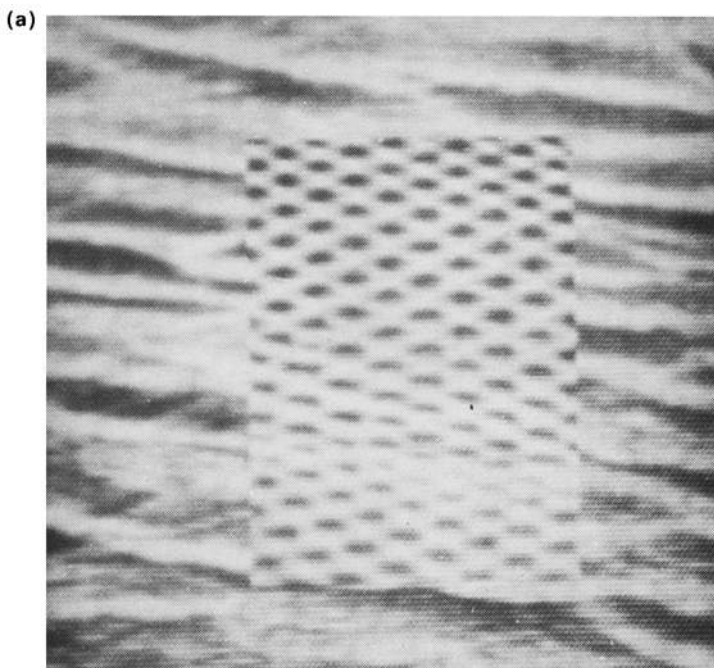


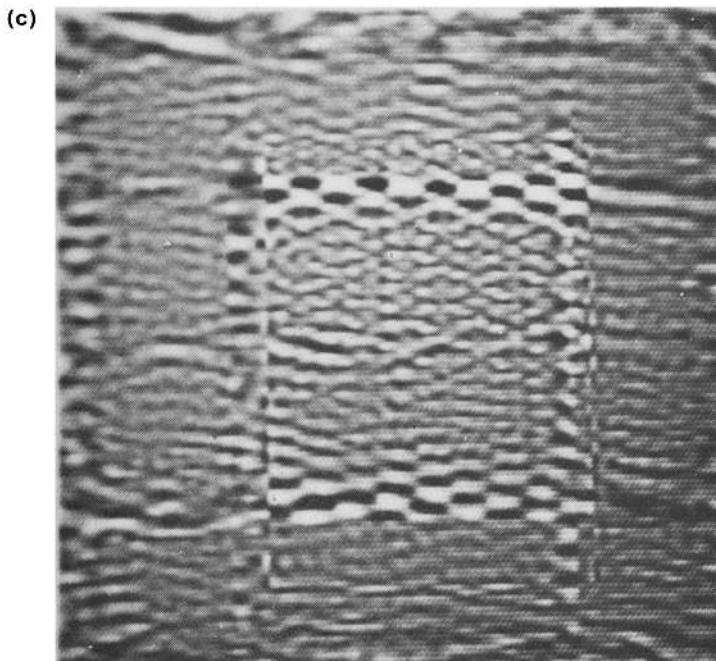
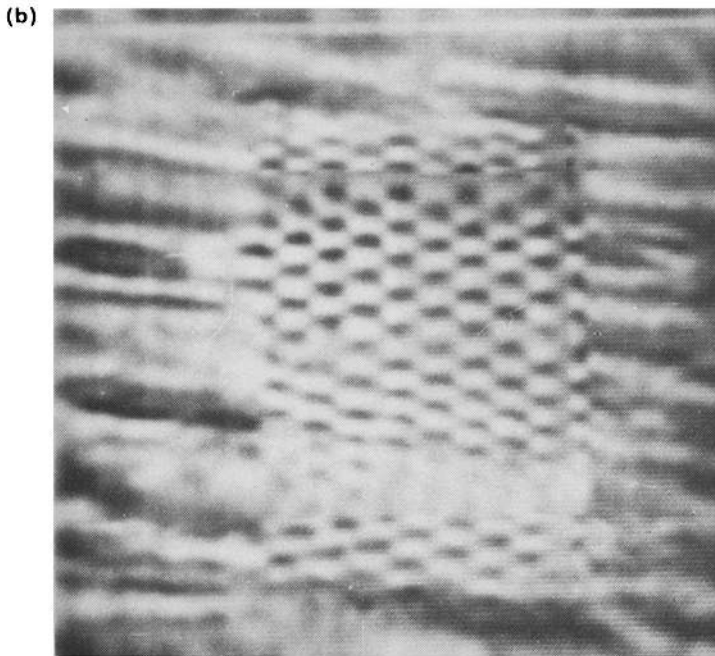
Plate 3



**Figure 3.** (a) Original image. (b) Reconstructed image (using 7 layers). (c) Residual error (at the end of 7 layers). Fourier transforms of layer 0 output (d), layer 1 output (e), layer 2 output (f), layer 3 output (g), layer 4 output (h), layer 5 output (i), and layer 6 output (j). Note that in figures (d) to (j), the outputs on the left hand side are real and those on the right are imaginary.



**Figure 4.** (a) (Caption on p. 324)



**Figure 4.** (a) Original image. (b) Reconstructed image (using 7 layers) and (c) residual error (at the end of 7 layers).