

# IMAGE REPRESENTATION USING VECTOR WAVELETS

Y.V.Venkatesh, K.Ramani and R.Nandini

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

## Abstract

*A new layered representation of images is proposed using, what one may call, vector wavelets defined by generalized Hermite polynomials. This representation has some special properties : (i) It is distinct from and superior to the other wavelet schemes of the literature ; (ii) It is stable ; (iii) It transforms the image into matrices of coefficients in the same manner as the standard transforms (Fourier, Hadamard and others), but at the specified 'scales' ; (iv) The zero-crossings of the signal at various scales can be directly obtained from the coefficients ; and (v) The size of the resolution cell in the 'phase-space' is variable even at a specified scale, depending on the image being analysed. This representation has been successfully applied to different types of images, both synthetic and natural.*

## 1 Introduction

It is known that biological (including human) vision employs multi-channel processing for low-level analysis of sensory data. One of the 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. Motivated by this discovery (Hubel and Wiesel [1]), 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.

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 suggested structure for implementing a multi-scale representation is called the pyramid [2]. Assuming a 2-D function  $f(x,y)$  defined over a digital grid,

one can define the pyramidal structure as a collection of subsampled images connected by a mapping transformation. A disadvantage of such a representation is that the data sets at separate levels are correlated.

### 1.1 Previous Work

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. A modification was made by employing a Gaussian window in an attempt to minimize the uncertainty associated with the spatial-spectral resolution, exemplified by the results of Marr [3], which involve the filtering of the original image with the Laplacian of a Gaussian for various values of the variance parameter. 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, finer details in an image when interspersed with coarse information cannot be separated out satisfactorily.

In order to overcome this deficiency of the window Fourier transform, a combined space-spatial frequency representation a la Gabor [4] or the so-called wavelet transform can be attempted. The former representation uses a modulated version of the Gaussian, but the Gabor functions do **not** constitute an orthogonal basis. 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 [5].

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 and Morlet [6] decompose a function in  $L^2(R)$  using the family of functions  $(\sqrt{s}\phi(sx)), s \in \mathcal{R}$ . 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. However, in practice, the standard procedure for computation is a decomposition of the image using the so-called "quadrature mirror" filters [7].

Consider a one-dimensional signal in  $L^2$ . Let  $\phi$

denote a function with sufficient decay, say  $\phi(x) \leq \frac{c}{(1+x^2)}$ , with

$$\int \phi dx = 0, \text{ and } \phi_s(x) = \frac{1}{s} \phi(x/s),$$

$$\phi_s^a = \frac{1}{s} \phi\left(\frac{x-a}{s}\right)$$

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

$$W(s, a) = \int f(x) \phi_s^a(x) dx$$

The choice of  $\phi$  determines the compactness of the representation. Inversion is achieved by an appropriate inverse integral. In practice, most commonly,  $s$  and  $a$  are restricted to some discrete subset:  $s = 2^{-m}$  and  $a = 2^{-n}$  with  $m, n \in \mathbb{Z}$ , generating a set of dyadic wavelets. The Haar basis is a standard example for discrete wavelets.

The generation 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 [7c] starts with an orthonormal basis of  $L^2(\mathbb{R})$  generated by the family of functions

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

However, there still remains the problem of extracting the coefficients of representation. It may be noted that orthogonality, apart from facilitating the computation of coefficients, guarantees their uniqueness [8b]. Other authors (see References in [7c]) have dealt with discrete wavelets  $\phi(x)$  with the property that  $(\sqrt{2^j} \phi(x))_{j \in \mathbb{Z}}$  constitute an orthonormal basis of  $L^2(\mathbb{R})$ . It is found that building up of such orthonormal bases of  $L^2(\mathbb{R})$  is quite involved.

## 2 Results based on a new vector wavelet transform

In practice, images are finite in spatial extent, and hence are not strictly finite in extent in the spectral domain. However, for mathematical analysis, they can be treated as being **approximately** infinite in both the spatial and spectral domains. To represent them, we employ generalized Hermite polynomials in multiple channels, each channel corresponding to the scale parameter  $\sigma$ . For each channel, the representation is in the explicit form of a matrix of coefficients.

### 2.1 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  $f(x, y) \in L^2(\mathbb{R})$  be a real-valued function of  $x, y \in \mathbb{R}$ , 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 [9][10].

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: Let

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

$$(x - x_0) = \Delta x_0, \quad (y - y_0) = \Delta y_0, \quad (\omega_1 - \omega_{10}) = \Delta \omega_{10}, \quad \text{and} \quad (\omega_2 - \omega_{20}) = \Delta \omega_{20}.$$

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

$$\sqrt{\frac{1}{E} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (\Delta x_0)^2 (\Delta y_0)^2 |f(x, y)|^2 dx dy}, \quad \text{and} \\ \sqrt{\frac{1}{4\pi^2 E} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (\Delta \omega_{10})^2 (\Delta \omega_{20})^2 |F(\omega_1, \omega_2)|^2 d\omega_1 d\omega_2}.$$

The corresponding inequality is:  $X_e W_e \geq 1$

### 2.2 Choice of Basis Functions and their properties

Consider the generalized 1-D Hermite polynomials parametrized by  $\sigma$ ,

$$P_n(z, \sigma) = \left( (-1)^n e^{\frac{z^2}{2\sigma^2}} \frac{d^n}{dz^n} e^{-\frac{z^2}{2\sigma^2}} \right)_{z = \frac{x}{\sigma}}$$

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  $\sigma_2$ ,

$$H_{m,n}(x, y, \sigma_1, \sigma_2) = P_m(x, \sigma_1) P_n(y, \sigma_2) \quad (2)$$

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

It is known [12][13] 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

For the 2-D basis representation, we use, in what follows, the vector notation,  $\underline{p}$ , 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 generalized Hermite polynomials can be shown to satisfy useful recurrence relations [11] and 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 have similar mathematical forms. 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  $\sigma_2$ , smaller the spatial width (and greater the spectral width), and vice versa, in the the directions  $x$  and  $y$  (and  $\omega_1$  and  $\omega_2$ ) respectively.

In what follows,  $\Sigma_{m,n}$  denotes summation with respect to  $m, n$  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  $m, n = 0, 1, 2, \dots$

Now we define formally the series,

$$f(x, y) = \Sigma_{m,n} \gamma_{m,n} H_{m,n}(\underline{p}), \quad -\infty < x, y < \infty \quad (4)$$

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

$$\gamma_{m,n} = \frac{1}{k_{m,n}} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) H_{m,n}(\underline{p}) dx dy \quad (5)$$

for  $m, n = 0, 1, 2, \dots$ . In practice, we use only a finite number,  $N$ , of terms. The coefficients  $\gamma_{m,n}$  are obtained by the standard procedure of minimizing the mean square error. As a consequence, the (approximate) image reconstructed from these coefficients will not match with the original at all the individual points. For convergence conditions, see [12],[13]. smooth in every finite interval  $[-a, a]$ , and if the integral

The error in the representation at scale  $(\sigma_{10}, \sigma_{20})$ , at any point  $(x, y)$ , is given by

$$err(x, y, \sigma_{10}, \sigma_{20}) = f(x, y) - f_{approx}(x, y), \quad (x, y) \in \mathcal{R}^2 \quad (6)$$

where the error is explicitly shown as dependent on  $\sigma_{10}$  and  $\sigma_{20}$ . Let this error be denoted by  $err_0$ . In fact, minimization of the Mean-square error is responsible for the inadequate representation of high frequency components in an infinite 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. As a consequence, an expansion of  $err$  using  $\sigma_1$  and  $\sigma_2$  smaller than  $\sigma_{10}$  and  $\sigma_{20}$  respectively can take into account the frequencies not found in  $f_{approx}$ . Let the error  $err(x, y, \sigma_{11}, \sigma_{21})$  in the representation at scale  $(\sigma_{11}, \sigma_{21})$ , at any point  $(x, y)$ , be denoted by  $err_1$ . Then

$$err_1 = err_0 - err_{approx}(x, y, \sigma_{11}, \sigma_{21}), \quad (x, y) \in \mathcal{R}^2. \quad (7)$$

By combining the above equations, it can be shown that

$$f(x, y) = f_{approx}(x, y) + err_{approx}(x, y, \sigma_{10}, \sigma_{20}) + err_{approx}(x, y, \sigma_{11}, \sigma_{21}) + \dots + \text{residual error} \quad (8)$$

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. By virtue of the multi-stage decomposition, the spectrum of  $f_{approx}(x, y)$  does not include that of  $err_{approx}(x, y, \sigma_{10}, \sigma_{20})$ , which in turn does not contain that of  $err_{approx}(x, y, \sigma_{11}, \sigma_{21})$ , and so on. This is equivalent to applying a sieve (the Hermite polynomials), at every level, that retain specific spectral information controlled by the scale parameter. The results of reconstruction are given in Fig.1.

### 3 Properties of the New Vector Wavelet Transform

There are many interesting properties of the new wavelet transform. However, in view of the constraints on space, details cannot given here. Refer to [11b]. For instance, zero-crossing contours in the image at various scales can be obtained directly from the coefficient matrices by exploiting the recursive relation among the Hermite polynomials. The other properties are : (i) The representation is stable with respect to coefficient perturbations ; (ii) The size of the phase-space resolution space is variable ; and (iii) The layered representations are independent.

### 4 Conclusions

An elegant method, using generalized Hermite polynomials and parametrized by a 'scale factor' (which leads to 'Vector Wavelets'), for representing images in terms of a matrix of coefficients has been proposed. The novelty of the results lies in the fact that the traditional assumption of finite support in the spatial or frequency domain has been dispensed with, while specifying an upper bound on the uncertainty in spatial-spectral localization.

### REFERENCES

1. D.H.Hubel and T.N.Wiesel, Receptive fields, binocular interaction and functional architecture in the cat's visual cortex, Journal of Physiology, Vol.160, pp 106-154, 1962.
2. P.J.Burt and E.H.Adelson, The Laplacian pyramid as a compact image code, IEEE Trans. Commun., Vol.COM-31, pp 532-540, April 1983.
3. D.Marr, Vision, San Francisco, CA : Freeman, 1982.
4. D.Gabor, Theory of Communication, J. Inst. Elec. Eng., (London), Vol.93 ,No.III, pp 429-457 , 1946.

5. M.J.Bastiaans, Gabor's signal expansion and degrees of freedom of a signal , Proceedings of the **IEEE** , Vol.68, pp 538-539, 1980.
6. A.Grossman and J.Morlet, Decomposition of Hardy functions into square integrable wavelets of constant shape, SIAM J.Math., Vol.15, pp 723-736, 1984.
7. S.Mallat, (a) Multi-resolution approximation and wavelet orthonormal bases of  $L_2$ , Trans. Amer. Math. Soc., Vol.3-15, pp 69-87, September 1989;  
 (b) A theory for multiresolution signal decomposition - the wavelet representation, Trans. PAMI-11, No. 7, pp 674-693, 1989;  
 (c) Multi-frequency channel decompositions of images and wavelet models, IEEE Trans.ASSP., Vol.37, No.12, pp 2091- 2110, 1989;
8. I.Daubechies, (a) The Wavelet Transform, time-frequency localization and signal analysis, IEEE Trans.IT, Vol.36, No.5, pp 961-1004, Sept. 1990;  
 (b) Orthonormal basis of compactly supported wavelets, Comm. Pure Applied Math., Vol. 41, pp 909-996, 1988.
9. D.L.Donoho and P.B.Stark, Uncertainty principles and signal recovery, SIAM Jl. Appl. Math., Vol.49, No.3, pp 906- 931, June 1989.
10. N.G.de Bruijn : Uncertainty Principles in Fourier Analysis in Inequalities, O.Shisha, Ed.,(Book) NewYork : Academic Press, pp 57-71, 1967.
11. Y.V.Venkatesh, K.Ramani and R.Nandini : (a) Vector wavelet decomposition of 1-D signals, Technical Report, Department of Elec. Engg. IISc, July 1991 ; (b) Hermite Sieve for Image Decomposition, Technical Report, Department of Elec. Engg. IISc, April 1992
12. N.N.Lebedev, Special Functions and Their Applications, Dover Publications, Inc. New York , pp 60-76, 1972.
13. J.R.Higgins, Completeness and Basis Properties of Sets of Special Functions, Cambridge, Cambridge University Press, Great Britain, 1977.
14. G.Szego, Orthogonal Polynomials, American Mathematical Society, 1975.

Fig. 1(a) Original Image



Fig.1(b) Reconstructed Image (7 Layered Output)



Fig.1(c) Residual Error Image (at the End of 7 Layers)

