The Multiscale Hermite Transform for Local Orientation Analysis

José L. Silván-Cárdenas and Boris Escalante-Ramírez

***Abstract*—The efficient representation of local differential structure at various resolutions has been a matter of great interest for adaptive image processing and computer vision tasks. In this paper, we derive a multiscale model to represent natural images based on the scale-space representation: a model that has an inspiration in the human visual system. We first derive the one-dimensional case and then extend the results to two and three dimensions. The operators obtained for analysis and synthesis stages are derivatives of the Gaussian smoothing kernel, so that, for the two-dimensional case, we can represent them either in a rotated coordinate system or in terms of directional derivatives. The method to perform the rotation is efficient because it is implemented by means of the application of the so-called generalized binomial filters. Such a family of discrete sequences fulfills a number of properties that allows estimating the local orientation for several image structures. We also define the discrete counterpart in which the coordinate normalization of the continuous case is approximated as a subsampling of the discrete domain.**

***Index Terms*—Hermite transform, local orientation, multiresolution decomposition, neighborhood operators, scale-space, steerable filters.**

I. Introduction

ONE OF THE earliest models of the human visual system, at cortex level, is based on the Gabor functions [1]–[4]. Filter design to simulate the visual processing carried out by some stages of the visual system has been of much concern in the field of visual processing information, such is the case of the cortex transform, in which analysis functions were designed to approximate the Gabor function profiles [5], with the advantage of being able to adjust both the radial and angular bandwidth. Many others have remarked on the need for a transform exhibiting a wide range of orientation and scale tuning characteristics for use with certain classes of objects common to all images [6]–[9].

In 1987, Young proposed a model based on the Gaussian and its derivatives [10]–[12]. Young showed that Gaussian derivatives fit more accurately to the measurements of the signal at the receptive fields than the Gabor function does, with an additional advantage of being orthogonal at the same location of analysis. Considering the spatial derivatives up to some order N enables characterization of the local image structure up to that order. In an early work [13], Koenderink and van Doorn advocated the use of this so-called *multiscale N-jet signal representation* as a model for earlier stages of visual processing. Then, in [14] they considered the problem of deriving linear operators from the scale-space representation [15]. They concluded that these operators must obey the time-independent Schrödinger equation, i.e., a physical equation that governs the quantum mechanical oscillator. Thus, they provided a formal statement that Gaussian derivatives are *natural operators* to derive from scale-space. These operators have been shown to be very effective for signal coding [16], [17] as well as useful for feature extraction [18], [19]. Some signal decomposition models based on these operators at single scale are described in [20]–[22]. In particular, the Hermite transform was originally introduced in [21] as a special case of a more general transform, namely, the polynomial transform, where some given requirements led to a Gaussian window function and the polynomial transform resulted in the Hermite transform. In that case, the Gaussian derivatives are interpreted as the Hermite polynomials multiplied by a Gaussian window. The signal within the window is expanded over the basis of Hermite polynomials and therefore called the Hermite transform. This transform was also inserted into a hierarchical structure for a multiscale setting in [23].

Freeman and Adelson developed a technique to steer filters by linearly combining basis filters oriented at a number of specific directions [6]. The Gaussian derivative family is perhaps the most common example of such functions. Since all Hermite filters are polynomials multiplied by a radially symmetric window function, it is easy to prove that the $n + 1$ Hermite filters of order n form a steerable basis for every individual filter of order n. Two distinct bases for the subspace of order n can be readily constructed: one from directional derivatives at sampled orientations [6] and the other from partial derivatives along orthogonal axes [19]. The possibilities are, in fact, infinite since the set of basis functions required to steer a function is not unique [24]. In a related work, Martens introduced several forms of the Hermite transform using a systematic approach developed mainly in the frequency domain and described their application for local orientation analysis [25].

In this paper, we present a similar transform based on constant octave-width frequency bands whose profiles resemble the ones found in biological visual systems [26], [27]. Early results of this work were reported in [28], [29] and [30]. In this paper, we explore in depth the theoretical results and discuss some practical considerations for certain applications. In Section II,

Manuscript received August 11, 2004; revised April 3, 2005. This work was supported by PAPIIT under Grants IN107101 and IN105505. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Ivan W. Selesnick.

J. L. Silván-Cárdenas was with the Centro de Investigación en Geogrfía y Geomática "Ing. J. L. Tamayo" (CentroGeo), Tlalpan, México, D.F. 14240 and also with the Department of Geography, Texas State University-San Marcos, San Marcos, TX 78666 USA (e-mail: jlsilvan@centrogeo.org.mx; js1536@txtstate.edu).

B. Escalante-Ramírez is with the Universidad Nacional Autónoma de México (UNAM), Facultad de Ingeniería, Cd. Universitaria, México, D.F. 04510 (e-mail: boris@servidor.unam.mx).

Digital Object Identifier 10.1109/TIP.2005.864177

we derive the one-dimensional (1-D) multiscale Hermite transform (MHT) which is, in contrast to the previously introduced Hermite transform, based on coefficients represented in the continuous spatial domain termed the *natural coordinates*. A natural coordinate system is defined by normalization of the spatial domain in proportion to the scale parameter of the representation. We also examine some properties of this signal decomposition from the point of view of adaptive signal processing and derive some pyramidal relations for its efficient implementation. Then, in Section III, we demonstrate how the two-dimensional (2-D) coefficients over a rotated coordinate system can be efficiently computed through a unitary transform based on the so-called generalized binomial family (GBF). An alternative representation based on directional Gaussian derivatives is also derived in the spatial domain and explicit relations to the 2-D separable case are provided. Applications to local orientation analysis have been the major concern in this work. It is well known that local orientation estimation can be achieved by combining the outputs from polar separable quadrature filters [9]. We illustrate here how the MHT can provide a mathematical insight for more detailed analysis of local structures. In Section IV, we present some remarks on the generalization of this decomposition to three-dimensional (3-D) signals. The discrete analogue of our signal decomposition is studied in Section V with some examples on natural images.

A. Notation

$\mathbb{Z}$, $\mathbb{R}$, and $\mathbb{C}$ denote the set of integer, real and complex numbers respectively. $\mathcal{L}^p(\mathbb{R})$ denotes the vector space of measurable, p-integrable 1-D real functions. The norm of $L(x) \in \mathcal{L}^p(\mathbb{R})$ is given by

$$(\|L\|_p)^p = \int_{-\infty}^{\infty} |L(x)|^p \, dx.$$

A p-integrable function satisfy $\|L\|_p < +\infty$.

For $J(x), L(x) \in \mathcal{L}^2(\mathbb{R})$, the inner product of $J(x)$ and $L(x)$ is written as

$$\langle J(x), L(x) \rangle = \int_{-\infty}^{\infty} J(x) L(x) dx.$$

The convolution of two functions $J(x)$ and $L(x) \in \mathcal{L}^2(\mathbb{R})$ is defined as

$$J(x) \star L(x) = \int_{-\infty}^{\infty} J(y) L(x-y) dy.$$

The identity element of $\mathcal{L}^2(\mathbb{R})$ under convolution is the *Dirac's Delta* function denoted by $\delta(x)$. Thus, $L(x) = \delta(x) \star L(x) = L(x) \star \delta(x)$.

The Fourier transform of $L(x) \in \mathcal{L}^2(\mathbb{R})$ is written $l(\omega)$ and is defined by

$$l(\omega) = \int_{-\infty}^{\infty} L(x) e^{-j\omega x} dx.$$

$\mathcal{L}^2(\mathbb{R}^2)$ is the vector space of measurable, square-integrable 2-D functions $L(x,y)$. For $J(x,y), L(x,y) \in \mathcal{L}^2(\mathbb{R}^2)$, the inner product of $J(x,y)$ and $L(x,y)$ is written as

$$\langle J(x,y), L(x,y) \rangle = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} J(x,y) L(x,y) dx dy.$$

We shall indicate the integral over a partial domain by using a subscript notation, e.g.,

$$\langle J(x,y), L(x,y) \rangle_x = \int_{-\infty}^{\infty} J(x,y) L(x,y) dx.$$

The Fourier transform of $L(x,y) \in \mathcal{L}^2(\mathbb{R}^2)$ is written $l(\omega_x, \omega_y)$ and it is defined by

$$l(\omega_x, \omega_y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} L(x,y) e^{-j(\omega_x x + \omega_y y)} dx dy.$$

The variables (u,v) and (ω_u, ω_v) are used for the rotated systems of spatial coordinates (x,y) and frequency coordinates (ω_x, ω_y) respectively. For a given angle θ these variables are related through

$$u = x\cos\theta + y\sin\theta, \quad v = -x\sin\theta + y\cos\theta$$
$$\omega_u = \omega_x\cos\theta + \omega_y\sin\theta, \quad \omega_v = -\omega_x\sin\theta + \omega_y\cos\theta.$$

$\mathcal{L}^2(\mathbb{Z})$ denotes the vector space of square-summable sequences $L(x)$ with $x \in \mathbb{Z}$. The norm in $\mathcal{L}^2(\mathbb{Z})$ is given by

$$\|L\|^2 = \sum_{x=-\infty}^{\infty} L^2(x).$$

The z-transform of $L(x) \in \mathcal{L}^2(\mathbb{Z})$ is written $l(z)$ and it is defined by

$$l(z) = \sum_{x=-\infty}^{\infty} L(x) z^{-x}$$

for $z \in \mathbb{C}$. This transform is related to the Discrete Fourier Transform through the change of variable $z = e^{j\omega}$.

The discrete convolution of $L(x)$ and $J(x) \in \mathcal{L}^2(\mathbb{Z})$ is defined by

$$J(x) \star L(x) = \sum_{y=-\infty}^{\infty} J(y) L(x-y).$$

The nth-order *forward difference* operator Δ^n is defined by

$$\Delta^n L(x) = \sum_{k=0}^{n} (-1)^{n-k} C_n^k L(x+k)$$

where

$$C_n^m = \binom{n}{m} = \frac{n!}{(n-m)!m!} = \frac{n(n-1)\cdots(n-m+1)}{m!}$$

denotes the binomial coefficients. This definition implicitly states $C_n^m = 0$ for $m < 0$ or $m > n$.

II. One-Dimensional Multiscale Hermite Transform

A. Multiscale Decomposition

A *scale-space representation* is a special type of multiscale representation that comprises a *continuous scale parameter* and preserves the same spatial sampling at all scales. The Gaussian scale-space, as introduced by Witkin [15], is an embedding of the original signal into a one-parameter family of derived signals constructed by convolution with Gaussian kernels of increasing width.

The scale-space representation of $L(x) \in \mathcal{L}^2(\mathbb{R})$ is denoted by $L(x;\sigma)$ for $\sigma \in \mathbb{R}^+$ and it is defined by letting the scale-space representation at zero scale be equal to the original signal $L(x;0) = L(x)$ and for $\sigma > 0$

$$L(x;\sigma) = L(x) \star G(x;\sigma)$$

where $G(x;\sigma)$ for $\sigma \in \mathbb{R}^+$, is the Gaussian kernel defined as the scaled Gaussian

$$G(x;\sigma) = \frac{1}{\sqrt{4\pi\sigma}} e^{-\frac{x^2}{4\sigma}}$$

normalized with respect to the $\mathcal{L}^1(\mathbb{R})$ norm. The normalized coordinate $x/\sqrt{4\sigma}$ is referred to as *natural coordinate* [14].

The scale-space theory generalizes the existing notions of Gaussian pyramids [31] and provides a well-founded way of relating image structure between different scales. It also provides a framework for early visual computations of a more general nature [13], [32]. Indeed, it has long been known that retinal ganglion cell receptive fields can be described using the difference of two Gaussian functions (DoG) of different scales (see, e.g., [26] and [27]). In [27], the retinal image is represented through a wavelet-like transform where the basic filters are DoG functions. Here, we use a similar idea as the starting point for the multiscale decomposition.

Formally, the impulse response of a DoG filter is defined by $DoG(x;\sigma;\alpha) = G(x;\sigma) - G(x;\alpha\sigma)$ for $\alpha > 1$. Thus, an input signal can be decomposed as

$$L(x) = \sum_{k=-\infty}^{\infty} L(x) \star DoG(x;\sigma_{k-1};\alpha) \tag{1}$$

where $\sigma_k = \alpha\sigma_{k-1} = \alpha^k\sigma_0$ for $k \in \mathbb{Z}$. The DoG filter captures the details that are to be added to a representation at scale $\alpha\sigma$ to obtain the representation at the lower scale σ. Unlike [27] here we use $\alpha = 4$ so the natural coordinates between levels maintain a factor of 2. Also, we fix $\sigma_0 = 0.5$ for normalization purposes.

B. Structural Analysis

Image analysis is a fundamental part of most image processing and computer vision tasks. The aim of image analysis is to derive features from visual data for further processing and analysis. In order to support recognition of several image structures embedded in the image data and to serve as a precursor to more detailed analysis (e.g., defining a local orientation or scale, or labeling a region as "corner," "edge," etc.) a local structural analysis is demanded. In such a case, derivative operators comprise a mathematical toolkit to discern among the different image structures [14], [33].

As the meaning "local" is dependent on the considered scale, the *intrinsic dimension* at a given position, i.e., the number of degrees of freedom required to describe the local structure (see, e.g., [9] and [34]) is scale dependent, too. In the scale-space representation, the local differential structure must be inferred from derivatives of the scaled representations of the signal. Fortunately, the nth-order derivative of $L(x;\sigma)$ can be obtained by convolving the input signal $L(x)$ with the Gaussian derivatives

$$G_n(x;\sigma) = \frac{d^n}{dx^n} G(x;\sigma) = \frac{(-1)^n}{\sqrt{(4\sigma)^n}} H_n\left(\frac{x}{\sqrt{4\sigma}}\right) G(x;\sigma) \tag{2}$$

where H_n denotes the nth-order Hermite polynomial for $n = 0, 1, 2, \ldots$ [35]. These Gaussian derivative functions obey scale-space properties that are better understood in the frequency domain. The Fourier transform of these functions is expressed in 1-D as

$$g_n(\omega;\sigma) = (j\omega)^n e^{-\omega^2\sigma}. \tag{3}$$

To simplify the notation the equivalences

$$g(\omega;\sigma) = g_0(\omega;\sigma); \quad g_n(\omega) = g_n\left(\omega, \frac{1}{4}\right)$$
$$g_n^*(\omega;\sigma) = \frac{g_n\left(\omega;\frac{\sigma}{2}\right)}{\sqrt{2^n n!}}; \quad g_n^*(\omega) = g_n^*\left(\omega;\frac{1}{2}\right)$$

and their spatial counterpart will be used in the sequel. Then, it must be clear that

$$g_n(\sqrt{4\sigma}\omega) = (\sqrt{4\sigma})^n g_n(\omega;\sigma)$$

Also, from the orthogonality of the Hermite polynomials [35], we have

$$\int_{-\infty}^{\infty} \frac{G_n^*(x) G_m^*(x)}{G^*(x)} dx = \delta_{n,m}. \tag{4}$$

Taking the mth derivative of (3) with respect to the scale parameter leads to the following identity expressed in the spatial domain:

$$\frac{\partial^m}{\partial\sigma^m} G_n(x;\sigma) = G_{2m+n}(x;\sigma) \tag{5}$$

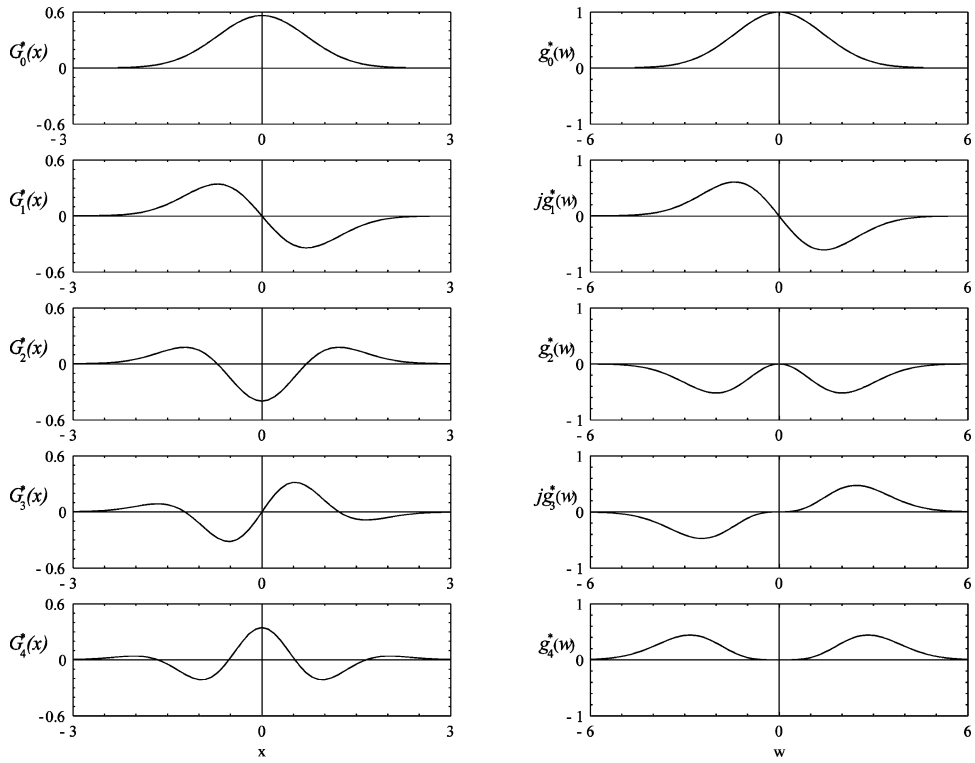

Fig. 1. Normalized Gaussian derivative filters in spatial and frequency domain for $n = 0, \ldots, 4$.

which is used, after setting $n = 0$, to obtain the Taylor expansion of $G(x; \sigma_{k-1})$ around σ_k, from which we subtracted the first term to write

$$DoG(x; \sigma_{k-1}; \alpha) = \sum_{n=1}^{\infty} \frac{(-\tau\sigma_k)^n}{n!} G_{2n}(x; \sigma_k) \tag{6}$$

with the *scaling constant*

$$\tau = \frac{\alpha - 1}{\alpha} = \frac{3}{4}.$$

In practice, we must limit the number of channels as well as the number of terms in the above expansion. There are indications from neurophysiologic studies that the human visual system works with derivatives up to a limited order [10]. In our initial model, however, we work with infinite expansions and, then, finite approximations can be treated as special cases.

C. Signal Representation

Adaptive image processing as well as image coding based on the signal decomposition approach generally requires a three-stage process. Within this approach the signal undergoes an *analysis* stage which simplifies the second stage of *processing* (or *coding*) and a third stage, the so-called *synthesis* stage, takes the intermediate representation to reconstruct a version of the original signal [25]. Under such a scheme, it is often advantageous that the synthesis stage performs similar operations to that of the analysis stage, i.e., the basis function to invert the transform are the same as the ones used to analyze the signal. This property is termed *self-invertibility* in the literature and has been recognized to be very important in many image processing applications [7]. Without this property, for instance, errors introduced by nonlinear processing of the coefficients (such as quantization) will spread to locations and frequencies other than those that were used to compute the coefficients.

Here, we factorized the even-order Gaussian derivatives in (6) using the well-known property of *concatenation* [14], i.e., $G_n(x; \sigma) \star G_m(x, \sigma') = G_{n+m}(x; \sigma + \sigma')$, to build a model that employs the same functions for both the analysis and the synthesis stages. Furthermore, if we express the convolutions as inner products we can write the signal decomposition of (1) as

$$L(x) = \sum_{k=-\infty}^{\infty} \sum_{n=1}^{\infty} \tau^n \left\langle \Lambda_n^{(k)}(\xi), 2^k \Psi_n^{(k)}(x; \xi) \right\rangle_{\xi} \tag{7}$$

where the signal is fully described in terms of the projections

$$\Lambda_n^{(k)}(\xi) = \left\langle L(x), \Psi_n^{(k)}(x; \xi) \right\rangle_x \tag{8}$$

of the input signal onto the basis functions

$$\Psi_n^{(k)}(x; \xi) = 2^{-k} G_n^*(2^{-k}x - \xi) \tag{9}$$

for $n \in \mathbb{N}$, $k \in \mathbb{Z}$, and $\xi \in \mathbb{R}$. Complying with the desired condition, these basis functions are used to both analyze and synthesize the signal and are nothing but the continuously shifted discrete-scaled version of the Gaussian derivatives normalized in the $\mathcal{L}^2(\mathbb{R})$ weighted norm. The functions for $n = 0, \ldots, 4$, $k = 0$ and $\xi = 0$, together with their Fourier transform, are plotted in Fig. 1.

One should notice that this base is nonorthogonal in the entire space of representation (n, k, ξ), but only with respect to the derivation order in the sense of (4). Fortunately, overcompletness is not of major concern for several image processing applications since orthogonality requires critical sampling constraint resulting in aliased representations that are unstable

under scaling and rotation of the input signal [7]. Our representation has some appealing properties that are discussed in the next sections.

D. Signal Details and Signal Smoothing

In order to determine what the above signal representation means and what kind of signal processing it can be applied to, we study the projections in (7) and (8) in the context of the scale-space representation.

We define the functions

$$L_n^{(k)}(x) = \left\langle \Lambda_n^{(k)}(\xi), 2^k \Psi_n^{(k)}(x;\xi) \right\rangle_\xi = \int_{-\infty}^{\infty} \Lambda_n^{(k)}(\xi) G_n^*(2^{-k}x - \xi) d\xi \tag{10}$$

for $n \in \mathbb{N}$, $k \in \mathbb{Z}$ and refer to them as the *derivative detail* functions since they provide the signal details of a given order n, at a given resolution k. In the Fourier domain, they are written as

$$l_n^{(k)}(\omega) = 2^k \lambda_n^{(k)}(2^k\omega) g_n^*(2^k\omega) \tag{11}$$

where $\lambda_n^{(k)}(\omega)$ is the Fourier representation of the multiscale Hermite transform (MHT) defined in (8), which is given by

$$\lambda_n^{(k)}(\omega) = 2^{-k} l(2^{-k}\omega) g_n^*(-\omega). \tag{12}$$

On the other hand, the set of smoothed signals $L^{(k)}(x) = L(x, \sigma_k)$ for $k \in \mathbb{Z}$ is called a *multiresolution approximation* of $L(x)$ because, at a given level of resolution, it contains all the details of coarser resolutions. Otherwise stated

$$L^{(k)}(x) = \sum_{j=k+1}^{\infty} \sum_{n=1}^{\infty} \tau^n L_n^{(j)}(x) = L^{(k+1)}(x) + \sum_{n=1}^{\infty} \tau^n L_n^{(k+1)}(x). \tag{13}$$

Notice that the second term in the righthand side of (13) corresponds to the signal information within the DoG channel with scales σ_k and σ_{k+1}. One can modify the width of this channel by simply altering the scaling constant τ, i.e., a particular way of processing the coefficients at level $k+1$. In particular, one may be interested in the resynthesis of the signal at an arbitrary scale based on the representation at discrete scales, i.e.,

$$L(x,\sigma) = L^{(k+1)}(x) + \sum_{n=1}^{\infty} (\tau\gamma)^n L_n^{(k+1)}(x) \tag{14}$$

where γ is referred to as the *smoothing factor* since for $\gamma = 0$ the signal equals $L(x, \sigma_{k+1})$ and, for $\gamma = 1$, it equals $L(x, \sigma_k)$. For other values of γ within 0 and 1 the signal equals the scaled representation at some scale σ between σ_k and σ_{k+1}. More precisely, if

$$\tau\gamma = \frac{\alpha' - 1}{\alpha'}$$

with $\sigma_{k+1} = \alpha'\sigma$, then $\sigma = (1-\gamma)\sigma_{k+1} + \gamma\sigma_k$ for $\gamma \in [0,1]$. Thus, the smoothing factor γ required to obtain the signal at some given scale σ through (14) can be easily found as

$$\gamma = \frac{4 - 2^{1-2k}\sigma}{3} \tag{15}$$

where the resolution level must be so that $\sigma_k \le \sigma \le \sigma_{k+1}$ and, therefore

$$k = \left\lfloor \log_2(\sqrt{2\sigma}) \right\rfloor \tag{16}$$

where $\lfloor x \rfloor$ denotes the largest integer lower than x.

E. Pyramidal Implementation

The processing scheme in which the image is decomposed into a number of bandpass or lowpass subimages, which are subsequently downsampled in proportion to their resolution, is called pyramidal representation, e.g., [31]. Each level in the pyramid maintains a linear relation to its adjacent levels and such relation is scale invariant. The hierarchical implementation of the Hermite transform [21] had been incorporated into a pyramid structure and applied successfully to noise reduction in computerized tomography imagery [36]. In that case, the pyramid scheme was used as an approximation of the multiscale decomposition, but no theoretical derivation of the decimation and interpolation filters was carried out. In this section, we show that the filter functions that generate a truly multiscale representations are completely specified in terms of the parameter α. With "truly" we mean that each level in the multiresolution representation has a dilation (scale) parameter associated in the corresponding basis function. In this respect, the pyramidal implementation is similar to that of a wavelet transform [37], [38].

The MHT can be implemented using this approach in the continuous spatial domain since from (12) we readily find

$$2\lambda_n^{(k+1)}(2\omega) = \lambda^{(k)}(\omega)\sqrt{\tau^{-n}} g_n^*(-\sqrt{3}\omega) \tag{17}$$

with $\tau = 3/4$, which can be written in the spatial domain as

$$\sqrt{\tau^n}\Lambda_n^{(k+1)}(\xi) = \int_{-\infty}^{\infty} \Lambda^{(k)}(\eta)\Phi_n(2\xi - \eta) d\eta \tag{18}$$

where

$$\Phi_n(x) = \frac{1}{\sqrt{3}} G_n^*\left(\frac{-x}{\sqrt{3}}\right) \tag{19}$$

is a Gaussian derivative filter of fixed scale. Equation (18) says that one only needs to filter the low-pass coefficient of the current level with some fixed scale functions, and scale down the output domain by a factor of two to generate the coefficients of the next resolution level.

On the other hand, if we write (13) in the frequency domain and use (11), we obtain the inverse relation

$$\lambda^{(k)}(\omega) = \sum_{n=0}^{\infty} \sqrt{\tau^n} 2\lambda_n^{(k+1)}(2\omega) g_n^*(\sqrt{3}\omega) \tag{20}$$

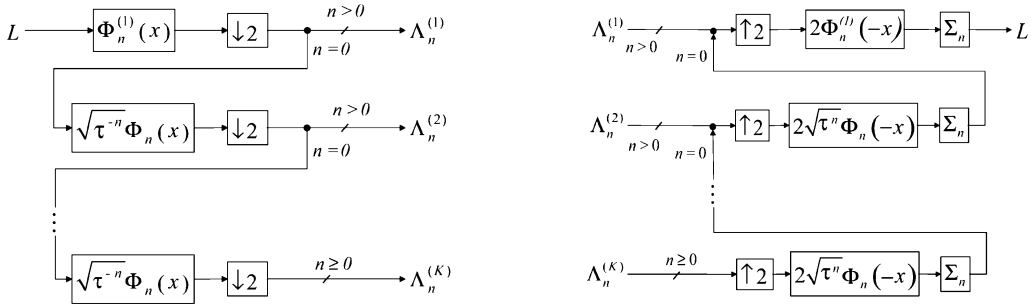

Fig. 2. Pyramidal implementation of analysis and synthesis stages for the multiscale Hermite transform. (a) Analysis. (b) Synthesis.

or, written in the spatial domain

$$\Lambda^{(k)}(\xi) = \sum_{n=0}^{\infty} \sqrt{\tau^n} \int_{-\infty}^{\infty} \Lambda_n^{(k+1)}(\eta) 2\Phi_n(2\eta - \xi) d\eta \qquad (21)$$

that is an *upscaling* relation through the filters $\Phi_n(x)$, for $n = 0, 1, 2, \ldots$ Therefore, the same set of filters allows mapping the low-pass component at a given resolution to or from the next lower resolution level with the advantage that they have a constant scale parameter. This certainly could make implementation more efficient; however, in order to make this pyramidal approach operable, we need to limit the number of scales of analysis.

If we consider a finite number K of channels, then from (13), we can write the multichannel decomposition as

$$L^{(1)}(x) = L^{(K)}(x) + \sum_{k=2}^{K} \sum_{n=1}^{\infty} \tau^n L_n^{(k)}(x) \qquad (22)$$

where σ_1 and σ_K determines the resolution limits of analysis. Although the low-pass residue $L^{(K)}(x)$ has lower variations than the input signal it can not be neglected for most practical situations. On the other hand, the high-pass residue $L(x) - L^{(1)}(x)$ can be neglected only if σ_1 is small enough compared to the information content of the signal. Unfortunately, we do not know *a priori* if relevant information is contained in the high-pass residue. Therefore, in order to recover the original signal, an additional channel from scale 0 to scale σ_1 (thus, with a unitary scaling constant) must be added to the above expansion. This results in the following signal expansion

$$L(x) = L^{(K)}(x) + \sum_{n=1}^{\infty} L_n^{(1)}(x) + \sum_{k=2}^{K} \sum_{n=1}^{\infty} \tau^n L_n^{(k)}(x) \qquad (23)$$

where the two first terms in the righthand side are precisely the low-pass and high-pass residues, respectively. The detail signals $L_n^{(1)}(x)$, for $n = 0, 1 \ldots$, are given in terms of the Hermite coefficients

$$\Lambda_n^{(1)}(\xi) = \int_{-\infty}^{\infty} L(x) \Phi_n^{(1)}(2\xi - x) dx \qquad (24)$$

where

$$\Phi_n^{(1)}(x) = \frac{1}{2} G_n^* \left(\frac{-x}{2} \right) \qquad (25)$$

for $n = 0, 1 \ldots$, the filters are required to generate the first layer of the pyramid. Subsequently, lower-resolution coefficients are generated through (18) for $k = 1, \ldots, K-1$. Conversely, the signal is reconstructed by integrating the highest resolution coefficients through

$$L(x) = \sum_{n=0}^{\infty} \int_{-\infty}^{\infty} \Lambda_n^{(1)}(\eta) 2\Phi_n^{(1)}(2\eta - x) d\eta. \qquad (26)$$

The complete analysis-synthesis process for the MHT is illustrated in the flow diagram of Fig. 2. All the filters required in the pyramidal scheme have the form of G_n^*, some of which are plotted in Fig. 1.

F. Predictive Scheme

An ideal representation of the local image should be based on a minimum number of parameters that, at the same time, are meaningful to aid interpretation. One obvious disadvantage of the MHT is the increment of the number of parameters to represent the input signal which gives a feeling of high redundancy for this representation.

Such redundancy is partially due to the linear dependency among analysis functions at different scales. Otherwise stated

$$\Lambda_n^{(k)}(\xi) = \sum_{m=0}^{\infty} b_{n,m}(\tau) \int_{-\infty}^{\infty} \Lambda_{m+n}^{(k+1)}(\eta) 2\Phi_m(2\eta - \xi) d\eta \qquad (27)$$

where $b_{n-m,m}(\tau) = \sqrt{C_n^m \tau^m (1-\tau)^{n-m}}$.

Notice that (27) is a generalization of (21) for higher order coefficients. Indeed, it can be further generalized to reconstruct the coefficients at any intermediate scale between σ_k and σ_{k+1} by simply changing the parameter τ as described in Section II-D. This property is termed *shiftability* in the scale due to Simoncelli *et al.* [7], a concept generalized from the *steerability* property first developed for the orientation domain by Freeman and Adelson [6]. Another example of scale-steerable family can be found in the literature [39], which seems to belong to other linear scale-space based on Poisson kernel [40].

In practice, one must limit the number of terms in this expansion when working on a discrete grid. One might think that the practical upper limit of derivatives will grow in proportion to the scale of the Gaussian. However, as scale increases, less details remain in the signal and then fewer derivatives are required. In

the discrete formulation of Section V, a fixed number of derivatives is used at each level with the exception of the first level. In such a case, we can only approximate the set of coefficients at a finer resolution using a partial sum similar to (27). This can be useful for coding applications where one typically predicts the coefficients of finer resolutions from those of coarser resolutions (which have lower entropy); then, one only transmits the coarsest level together with the prediction errors of finer levels. The mechanism is analogous to oversampling the signal and coding the differences between contiguous samples instead of coding the samples themselves since an increase in the sampling rate reduces the total data entropy.

III. Two-Dimensional Multiscale Hermite Transform

Most of the properties of the Gaussian derivatives in 1-D are readily extended to higher dimensions, since the multivariate Gaussian derivatives are defined as the product of 1-D Gaussian derivatives. In particular, the 2-D Gaussian derivatives are written in the frequency domain as $g_{m,n-m}(\omega_x,\omega_y;\sigma) = g_m(\omega_x;\sigma)g_{n-m}(\omega_y;\sigma)$, for $m = 0,\ldots,n$; and $n = 0,2,\ldots$; where m is the derivative order with respect to x, $n-m$ is the derivative order with respect to y, and n is the total derivative order. Then, property (5) is generalized to 2-D as

$$\frac{\partial^m}{\partial\sigma^m}G_{i,n-i}(\cdot;\sigma) = \sum_{j=0}^{m} C_m^j G_{2j+i,2m+n-2j-i}(\cdot;\sigma) \tag{28}$$

where $(\cdot)$ implies the spatial coordinates (x,y). Therefore, if $i = n = 0$, we can write the Taylor expansion for the 2-D DoG function as

$$DoG(\cdot;\sigma_{k-1};\alpha) = \sum_{n=1}^{\infty}\sum_{m=0}^{n}\frac{(-\tau\sigma_k)^n}{(n-m)!m!}G_{2m,2n-2m}(\cdot;\sigma_k). \tag{29}$$

Then, proceeding as in the 1-D case, we find the following expressions for the 2-D signal expansion

$$L(x,y) = \sum_{k=-\infty}^{\infty}\sum_{n=1}^{\infty}\sum_{m=0}^{n}\tau^n L_{n-m,m}^{(k)}(x,y) \tag{30}$$

where the detail signals

$$L_{n-m,m}^{(k)}(\cdot) = \left\langle \Lambda_{n-m,m}^{(k)}(\xi,\eta), 4^k\Psi_{n-m,m}^{(k)}(\cdot;\xi,\eta)\right\rangle_{(\xi,\eta)} \tag{31}$$

are based on the projections

$$\Lambda_{n-m,m}^{(k)}(\xi,\eta) = \left\langle L(\cdot), \Psi_{n-m,m}^{(k)}(\cdot;\xi,\eta)\right\rangle_{(x,y)} \tag{32}$$

of the signal onto the 2-D basis functions

$$\Psi_{n-m,m}^{(k)}(x,y;\xi,\eta) = 4^{-k}G_{n-m,m}^{*}(2^{-k}x-\xi, 2^{-k}y-\eta) \tag{33}$$

for $m = 0,\ldots,n$; $n \in \mathbb{N}$; $k \in \mathbb{Z}$, and $(\xi,\eta) \in \mathbb{R}^2$. These functions are separable in x and y, since $G_{n,m}^{*}(x,y) = G_n^{*}(x)G_m^{*}(y)$; therefore the 2-D MHT can be implemented by two cascaded 1-D MHT, each over one spatial dimension.

Besides the scale-space properties, the Gaussian derivatives in two or more dimensions exhibit many other symmetry properties related to *rotation* of the coordinate system that are discussed in the next section.

A. Rotated MHT

The first step in our signal analysis consists of the determination of the Cartesian Hermite transform, since associated filters are separable and maintain the pyramidal relations derived in Section II-E and, therefore, they can be efficiently implemented. For image analysis purposes, however, further steps are required. Martens [25] has suggested mapping the Cartesian coefficients to polar coefficients since finding the axis of symmetry and adaptively rotating the coordinate axes is done most easily on these polar coefficients. However the adaptive polar Hermite coefficients are generally mapped back to the Cartesian representation for interpretation, coding or processing. The approach presented in this paper differs from Martens approach in that the adaptive rotation is performed directly on the Cartesian coefficients with the additional advantage that the angle functions required to steer the 2-D isotropic Gaussian derivatives can be efficiently implemented.

Let $g_{n-m,m}(\omega_x,\omega_y;\sigma;\theta)$ denote the clockwise-rotated version of $g_{n-m,m}(\omega_x,\omega_y;\sigma)$ by an angle θ, i.e.,[1] $g(\omega_x,\omega_y;\sigma;\theta) = g(\omega_u,\omega_v;\sigma)$.

An explicit expression for the above rotated functions is readily obtained in the frequency domain (e.g., see [19], [29]) and written in the spatial domain as

$$G_{n-m,m}(\cdot;\sigma;\theta) = \sum_{k=0}^{n} A_m(k;n;\theta)G_{k,n-k}(\cdot;\sigma) \tag{34}$$

where the angular functions

$$A_n(k;N;\theta) = s^k c^{-k}\Delta^n\left\{C_{N-n}^{k-n}c^{2k-n}s^{N-2k+n}\right\} \tag{35}$$

for $k,n = 0,\ldots,N$; $N \in \mathbb{N}$; and $c = \cos\theta$, $s = \sin\theta$, for $\theta \in [0,2\pi)$, are referred to as the *generalized binomial family* (GBF) of index k. The reason for this name becomes clear after noticing that the filter transfer function in the z-transform domain (see Appendix I)

$$a_n(z;N;\theta) = (c - z^{-1}s)^n(s + z^{-1}c)^{N-n} \tag{36}$$

can be implemented by cascading the binomial kernels $\{c,-s\}$ and $\{s,c\}$, by n and $N-n$ times, respectively. An efficient implementation of these filters can be achieved, for instance, by generalizing the algorithm of [41] to include nonunitary weights given by these kernels. In this case, the number of operations per sample will be multiplied by three.

The GBF satisfies a number of properties that allow addressing problems concerning local orientation analysis within a mathematical framework. An exhaustive study of the properties of such functions is beyond the scope of this paper. However, we will cover some interesting mathematical relations of these sequences and provide an indication to derive them. Some of the properties are directly derived from the definition. For instance, the symmetries

$$(-1)^n A_n(N-k;N;\theta) = (-1)^k A_{N-n}(k;N;\theta) \tag{37a}$$

$$C_N^n A_n(k;N;\theta) = C_N^k A_k(n;N;\theta) \tag{37b}$$

$$A_n(k;N;-\theta) = A_{N-n}(N-k;N;\theta) \tag{37c}$$

[1]Thus, $g_{n-m,m}(\cdot;\sigma)$ and $g_{n-m,m}(\cdot;\sigma;0)$ can be indistinctly used.

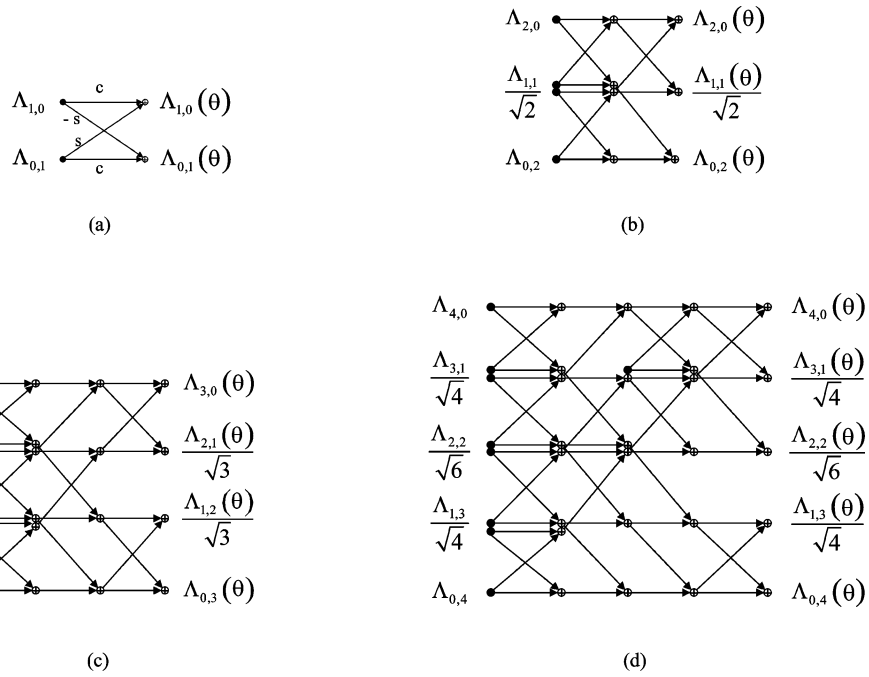

Fig. 3. Flow diagrams for rotating the Hermite transform coefficients up to fourth order. Horizontal arrows are weighted with $c = \cos\theta$, upward and downward arrows are weighted with s and $-s$, respectively, with $s = \sin\theta$. (a) $n = 1$; (b) $n = 2$; (c) $n = 3$; and (d) $n = 4$.

are easily verified from (35). But some other relations will require further development.

The original Gaussian derivatives are obtained from the rotated ones by applying the same filters with a change of sign of the angle parameter. This results in

$$G_{m,n-m}(\cdot;\sigma) = \sum_{k=0}^{n} A_m(k;n;\theta) G_{n-k,k}(\cdot;\sigma;\theta). \tag{38}$$

Equations (34) and (38) are direct consequences of the steering property satisfied by derivative filters [6], [19]. These linear mappings define a unitary transformation carried out over the partial derivation order of the 2-D Gaussian derivatives through the base comprised by the functions[2] $A_m(k;n;\theta)$, for $k, m = 0,\ldots,n$.

The expressions for the forward and backward rotation of the Hermite coefficients are found after normalizing and applying the Gaussian derivatives in the above expressions to the input signal. In order to simplify the notation let $\Lambda_{n-m,m}$, for $m = 0,\ldots,n$ denote the Hermite coefficients of order n at some generic location and scale, and let $\Lambda_{n-m,m}(\theta)$ denote the same coefficients over rotated axes by a given angle θ measured clockwise with respect to axis x. Then, from (34), the following relation for these coefficients is found:

$$\frac{\Lambda_{n-m,m}(\theta)}{\sqrt{C_n^m}} = \sum_{k=0}^{n} A_m(k;n;\theta) \frac{\Lambda_{k,n-k}}{\sqrt{C_n^k}}. \tag{39}$$

Furthermore, since the GBF for the difference of two angles can be evaluated by concatenation of the filters with the given angles (see Appendix II), a more general expression for the rotated coefficients can be written as

$$\frac{\Lambda_{n-j,j}(\theta-\phi)}{\sqrt{C_n^j}} = \sum_{m=0}^{n} A_{n-j}(m;n,\phi) \frac{\Lambda_{n-m,m}(\theta)}{\sqrt{C_n^m}} \tag{40}$$

from which the original coefficients are obtained by letting $\phi = \theta$.

Therefore, the rotation algorithm is rather simple. We first de-normalize the coefficients by dividing by a factor $\sqrt{C_n^k}$, then apply the GBF in cascaded fashion and, finally, multiply again by a factor $\sqrt{C_n^m}$. The flow diagrams for rotating the coefficients up to fourth order are shown in Fig. 3. As it can be inferred from the basic butterfly diagram of Fig. 3(a), there are twice as many multiplications as additions. Meanwhile, a total of $\lceil (n+1)[3n - \log_2(n+1)]/2 \rceil$ additions are required for the rotation of the nth-order coefficients.

B. Anisotropic Smoothing

The smoothing method described in Section II-D is easily generalized to two dimensions with independent smoothing factors for each dimension which, combined with the local rotation process, allows an efficient anisotropic Gaussian smoothing of the image.

We process the rotated coefficients to generate the *smoothed coefficients* through

$$\frac{\tilde{\Lambda}_{j,n-j}}{\sqrt{C_n^j}} = \sum_{m=0}^{n} A_j(m;n;\theta) \frac{\gamma_u^{n-m}\gamma_v^m \Lambda_{n-m,m}(\theta)}{\sqrt{C_n^m}} \tag{41}$$

where γ_u and $\gamma_v \in [0,1]$ denote the smoothing factors along directions u and v, the rotated coordinates that make an angle θ with respect to the original coordinates x and y respectively.

[2]This base is orthonormal in the sense of (72) in the Appendix II.

The smoothed coefficients $\tilde{\Lambda}_{j,n-j}$, for $j = 0, \ldots, n$, are still computed through the generalized binomial family since the filter $\gamma_u^{n-k}\gamma_v^k A_m(k; n; \theta)$ has a transfer function

$$a_m(z; n; \theta; \gamma_u, \gamma_v) = (\gamma_u c - z^{-1}\gamma_v s)^m (\gamma_u s + z^{-1}\gamma_v c)^{n-m} \tag{42}$$

which defines a more general form of the GBF.

The image is then adaptively smoothed by processing the coefficients at different positions independently and by choosing the smoothing parameters for the kth layer according to

$$\gamma_{u|v} = \begin{cases} 0, & \text{if } k < j \\ \frac{4-2^{1-2k}\sigma_{u|v}}{3}, & \text{if } k = j \\ 1, & \text{if } k > j \end{cases} \tag{43}$$

with $j = \lfloor \log_2(\sqrt{2\sigma_{u|v}}) \rfloor$, where $\gamma_{u|v}$ stands for either γ_u or γ_v and so $\sigma_{u|v}$ for the scales σ_u or σ_v. The scales $\sigma_{u|v} \in [0, \sigma_K]$ and the orientation $\theta \in [0, \pi)$ are free parameters. The natural application of this anisotropic Gaussian smoothing is to noise reduction, a problem which has been proven to be very hard to solve, mainly because one usually needs to estimate the orientation parameter from the noisy data (cf. [42]). One obvious limitation for this approach to noise reduction is its inability to eliminate noise around boundaries with high curvature (e.g., corner and junctions).

In the next subsection, we develop a more detailed structural analysis in the domain of the rotated MHT.

C. Local Orientation Analysis

The *intrinsic* dimension of a local structure in the image is the number of degrees of freedom required to describe it [34]. We shall say that an image is locally i-D if the intrinsic dimensions is $i = 0$, 1, or 2. For a locally 1-D image with normal along θ_0, some of the rotated coefficients vanish at that orientation, i.e., $\Lambda_{n-m,m}(\theta_0) = \Lambda_n \delta_{m,0}$, where Λ_n, for $n = 0, 1, \ldots$, are the 1-D Hermite coefficients of the image profile. Whereas for an arbitrary rotation they have the following angular variation

$$\Lambda_{n-m,m}(\theta) = \sqrt{C_n^m} A_m(n; n; \theta - \theta_0)\Lambda_n.$$

Then, the original 2-D coefficients

$$\Lambda_{k,n-k} = \frac{A_0(k; n; \theta_0)}{\sqrt{C_n^k}}\Lambda_n \tag{44}$$

can be used to compute the local orientation θ_0 since, using the trigonometric relations of Appendix III, we can write

$$\tan(n\theta_0) = \frac{\sum_{k=0}^{\lfloor \frac{n-1}{2} \rfloor} (-1)^k \sqrt{C_n^{2k+1}} \Lambda_{n-2k-1,2k+1}}{\sum_{k=0}^{\lfloor \frac{n}{2} \rfloor} (-1)^k \sqrt{C_n^{2k}} \Lambda_{n-2k,2k}}. \tag{45}$$

It must be clear that (45) leads to n solutions for θ_0 in the interval $[0, \pi)$ given in terms of the nth-order coefficients and they are uniformly spaced by π/n. For the 1-D patterns, we must expect one of the obtained solutions for $n > 1$ equals the gradient angle, i.e, the single solution for $n = 1$. Fig. 4 shows some examples of different patterns. The orientations obtained from (45) with $n = 1$ through 4 are shown as compass plots. Pattern labeled "A" is strongly oriented and, consequently, one of the solutions for each $n > 1$ coincides with the gradient orientation. In order to pick up the true normal orientation for $n > 1$ one can determine which of the n solutions gives the maximum energy for the directional derivative $\Lambda_{n,0}(\theta_0)$.

The 1-D patterns belong to a more general set of image structures namely the symmetric patterns, which includes 2-D structures, such as corners and symmetric junctions. For these patterns the local *axis of symmetry*, commonly used in computer vision for object recognition, can be directly determined from the Hermite coefficients. Indeed, (45) gives the orientation of the axis of symmetry. In order to show this, we use the fact that rotated coefficients of a mirror-symmetric pattern around the angle θ_0 satisfy $\Lambda_{n-m,m}(\theta_0) = u0$, for m odd [25]. Then, considering (40) with $\phi = \theta = \theta_0$ and the symmetry property given in (37b), we can write the unrotated 2-D coefficients as

$$\sqrt{C_n^j}\Lambda_{n-j,j} = \sum_{m=0}^{\lfloor \frac{n}{2} \rfloor} \sqrt{C_n^{2m}} A_{2m}(n-j; n, \theta_0)\Lambda_{n-2m,2m}(\theta_0). \tag{46}$$

The next steps consist in taking the cases $j = 2k$ and $j = 2k+1$, multiply them by $(-1)^k$, sum over k, and use the identities (73) and (74) derived in the Appendix III. The resulting expressions are finally combined by division to obtain (45). The edge normal orientation is, therefore, the symmetry axis for the particular case of 1-D patterns. As an example consider the pattern labeled "B" in Fig. 4. In this case, the angles selected as described before are all around 60°. This indicates that the pattern is approximately symmetric around that orientation. Unlike the edge of sample "A," the pattern of "B" correlates better with the second-order derivatives than with the first-order ones. The orientations produced by the former clearly define the two axes of symmetry observed in the image. The pattern labeled "C" is also roughly symmetric but it only has one axis of symmetry around 100°. The first- and second-order operators most likely describe this structure as they have higher output. In contrast to symmetric patterns, nonsymmetric patterns (e.g., sample "D" in Fig. 4) produce very different orientations for different derivation orders.

Interestingly, (45) gives also the solution to the general gauge condition [33]

$$\sum_{k=0}^{\lfloor \frac{n-1}{2} \rfloor} (-1)^k \sqrt{C_n^{2k+1}} \Lambda_{n-2k-1,2k+1}(\theta_0) = 0 \tag{47}$$

with rotated coefficients of not necessarily symmetric patterns. This can be verified by substituting (39) of the rotated coefficients into (47) and using the identities of Appendix III to obtain (45). Similarly, one can also derive the identity

$$\begin{aligned} &\sum_{k=0}^{\lfloor \frac{n}{2} \rfloor} (-1)^k \sqrt{C_n^{2k}} \Lambda_{n-2k,2k}(\theta_0) \\ &\quad = \left\{ \left[\sum_{m=0}^{\lfloor \frac{n}{2} \rfloor} (-1)^m \sqrt{C_n^{2m}} \Lambda_{n-2m,2m} \right]^2 \right. \\ &\qquad \left. + \left[\sum_{m=0}^{\lfloor \frac{n-1}{2} \rfloor} (-1)^m \sqrt{C_n^{2m+1}} \Lambda_{n-2m-1,2m+1} \right]^2 \right\}^{\frac{1}{2}} \end{aligned} \tag{48}$$

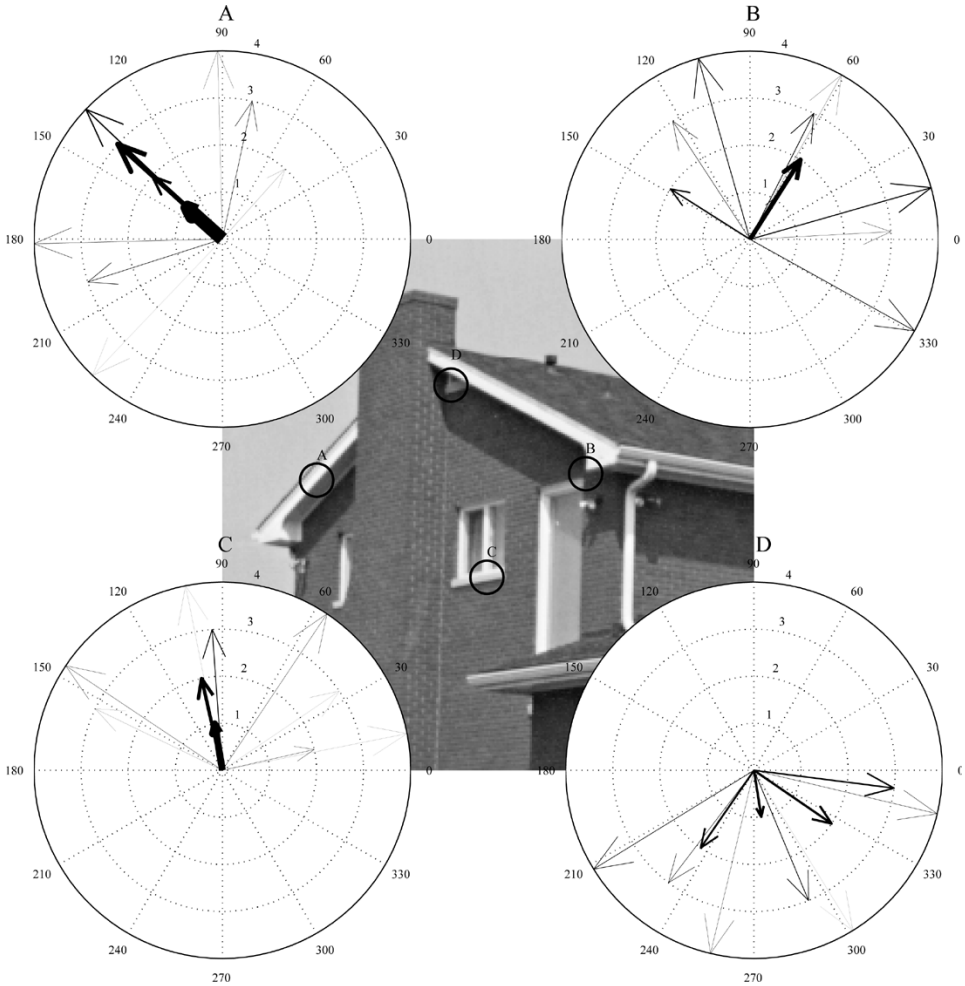

Fig. 4. Orientation analysis of different patterns. The compass plots show the orientations given by $\theta = \arctan(Q)/n + \pi k/n$ for $k = 0, 1 \ldots, n-1$, where Q is the right-hand term of (45) computed with the coefficients at the first resolution level of the discrete MHT. The length of each arrow was set to the derivation order n = 1,2,3,4 and its width adjusted proportionally to the magnitude of $\Lambda_{n,0}(\theta)$. Whenever $\Lambda_{1,0}(\theta)$ had a negative value, $\theta + \pi$ was taken to ensure the coherence among the estimation of different orders.

which defines a positive image measure that is invariant under rotation.

Therefore, for a given n, (47) can be regarded as a general criterion to fix the rotation angle regardless of the shape of the local pattern. In particular, the first gauge condition $\Lambda_{0,1}(\theta_0) = 0$ leads to the gradient orientation and reduces (48) to the gradient magnitude. The case $n = 2$ leads to the orientation of the eigenvectors for the Hessian matrix, i.e., it corresponds to the diagonalization condition for the Hessian matrix since (47) becomes $\Lambda_{1,1}(\theta_0) = 0$.

In many applications, it is useful to classify the local structure into a finite set of classes in order to both process each kind of structure more efficiently and facilitate further analysis and interpretation. Such classification should take place within the appropriate energy space (e.g., [29] and [43]). In Fig. 5, we illustrate the space covered by certain patterns within an energy space defined by the measured 1-D energy E^{1D} (i.e., the energy of the rotated coefficients with $n > 0$ that are nonzero for an ideal 1-D pattern), symmetric energy E^{SYM} (i.e., the energy of the rotated coefficients with $n > 0$ that are nonzero for an ideal symmetric pattern) and 2-D energy E^{2D} (i.e., the energy of all the coefficients with $n > 0$). Thus, any image pattern is represented by a point within the space limited by four planes. Each subspace outlined in this graphic represents a class of patterns grouped according to their dimensionality and symmetry. Flat patterns are located at the origin of the space, 1-D patterns are located along an edge of the pyramid and so forth. Unfortunately, real images are generally corrupted by noise. Since errors are

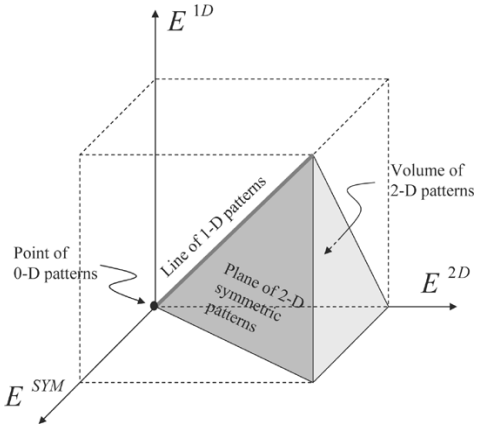

Fig. 5. Localization of different patterns within an energy space.

necessarily introduced due to approximations of the processing model and/or the arithmetic precision of the processing system, and because a degree of subjectiveness is involved in the separation of the above patterns, the classification problem is more complex in practice. A more realistic approach should allow the spaces defined by the different patterns to fuzzily overlap each other.

D. Multidirectional MHT

Freeman and Adelson introduced a technique to synthesize filters of arbitrary orientation from linear combination of basis filters chosen as rotated replicas of the so-called *steerable* function [6]. Directional Gaussian derivatives are perhaps the most common example of such functions. In a related work, Martens introduced a form of the Hermite transform termed the *Wavelet Hermite transform* based on directional derivatives [25]. In that case, the transform was defined through a systematic approach developed in the frequency domain using the fact that the Fourier transform of Gaussian derivative filters are separable and that the angular functions comprise a base for the Hilbert space of periodic signals. Here, we use properties of the GBF to derive a multidirectional form for the MHT. We also provide explicit relations between the 2-D separable MHT and the multidirectional MHT.

The generalized binomial sequences, when seen as angular functions, can be expanded into their Fourier series with the Fourier coefficients given in terms of the symmetric filters ((75) in Appendix III). From (75b), we notice that the DC coefficient c_0, i.e., the mean value of the angular function, is zero if either k, n, or N are odd, and for even values, we can write the integral expression

$$\frac{1}{2\pi}\int_0^{2\pi} A_{2n}(2k;2N;\theta)d\theta = 2^{-2N}C_{2N}^{N}C_{N}^{k}\frac{C_{N}^{n}}{C_{2N}^{2n}}.$$

Moreover, from some well-known properties of periodic signals, it can also be written as the discrete sum

$$\frac{1}{N+1}\sum_{j=0}^{N} A_{2n}(2k;2N;\theta_j) = 2^{-2N}C_{2N}^{N}C_{N}^{k}\frac{C_{N}^{n}}{C_{2N}^{2n}} \tag{49}$$

for $\theta_j = \theta_0 + j\pi/(N+1)$. This expression is used to relate the *directional* derivatives at uniformly spaced orientations to the *Cartesian* derivatives of the Gaussian kernel. Specifically, we set $n = 0$ in (49) and combine with (34) to obtain the identity

$$r_n\sum_{j=0}^{n} G_{2n,0}(\cdot;\sigma;\theta_j) = \sum_{m=0}^{n} C_{n}^{m}G_{2m,2n-2m}(\cdot;\sigma) \tag{50}$$

with

$$r_n = \frac{2^{2n}}{(n+1)C_{2n}^{n}}$$

for $n \in \mathbb{N}$. Therefore, from (29), we can write an alternative expression for the Taylor expansion of the DoG as

$$DoG(\cdot;\sigma_{k-1};\alpha) = \sum_{n=1}^{\infty}\sum_{j=0}^{n}\frac{(-\tau\sigma_k)^n}{n!}r_n G_{2n,0}(\cdot;\sigma_k;\theta_j)$$

in terms of directional derivatives at orientations $\theta_j = \theta_0 + j\pi/(n+1)$ for $j = 0,\ldots,n$. This expansion, in turn, defines another Hermite transform so that the input signal is expressed as

$$L(x,y) = \sum_{k=-\infty}^{\infty}\sum_{n=1}^{\infty}\sum_{j=0}^{n} r_n\tau^n L_n^{(k,j)}(x,y) \tag{51}$$

where the detail signals

$$L_n^{(k,j)}(\cdot) = \left\langle \Lambda_n^{(k,j)}(\xi,\eta), \Psi_n^{(k,j)}(\cdot;\xi,\eta)\right\rangle_{(\xi,\eta)} \tag{52}$$

are based on the projections

$$\Lambda_n^{(k,j)}(\xi,\eta) = \left\langle L(x,y), \Psi_n^{(k,j)}(x,y;\xi,\eta)\right\rangle_{(x,y)} \tag{53}$$

of the signal onto the 2-D basis functions

$$\Psi_n^{(k,j)}(x,y;\xi,\eta) = 2^{-k}G_{n,0}^{*}(2^{-k}u-\xi, 2^{-k}v-\eta) \tag{54}$$

for $n \in \mathbb{N}; k \in \mathbb{Z}; j = 0,\ldots,n; (\xi,\eta) \in \mathbb{R}^2$ and (u,v) is the rotate coordinate system. These basis functions are nonseparable in x and y. However, in practice, we can compute the multidirectional coefficients through the separable ones since

$$\Lambda_n^{(k,j)} = \sum_{m=0}^{n} A_0(m;n;\theta_j)\frac{\Lambda_{m,n-m}^{(k)}}{\sqrt{C_n^m}} \tag{55}$$

holds for a generic location. Moreover, it can be shown from (50) that the signal reconstruction is also achieved by applying the inverse MHT to the coefficients

$$\hat{\Lambda}_{m,n-m}^{(k)} = r_n\sum_{j=0}^{n} A_0(m;n;\theta_j)\frac{\Lambda_n^{(k,j)}}{\sqrt{C_n^m}} \tag{56}$$

for $m = 0,\ldots,n$; $n \in \mathbb{N}$ and $k \in \mathbb{Z}$. Notice that the sets $\{\hat{\Lambda}_{m,n-m}^{(k)}\}_{m=0,\ldots,n;n\in\mathbb{N}}$ and $\{\Lambda_{m,n-m}^{(k)}\}_{m=0,\ldots,n;n\in\mathbb{N}}$ are different; however, they synthesize the same image. This is only explained by the fact that the coefficients at different locations are not independent at all.

IV. Three-Dimensional Multiscale Hermite Transform

Although the most frequently used images are in two dimensions, 3-D datasets often arise in medical imagery as well as in video sequences. In this last case, a spatiotemporal signal is mapped onto a 3-D spatial signal by transforming the temporal coordinate to a spatial coordinate via multiplication by a velocity parameter u that is implicity selected with the sampling rate [21].

The definition of the 3-D MHT is straightforward. All the expressions for this case have additional indexes for the spatial coordinate and for the derivative order along the third dimension. Here, we restrict ourselves to some remarks for the rotation of the 3-D coefficients denoted by $\Lambda_{l,m,n}$ for $l,m,n \in \mathbb{N}$.

Rotation makes sense only if the resultant coordinates have a geometrical meaning regarding the local pattern of the image. For 3-D imagery, we must specify the transformation in terms of the kind of patterns we want to characterize. For instance, we can easily characterize 2-D patterns underlying constant movement along linear trajectories (at least locally and within a short period of time) if we perform the following two rotations. First,

perform a rotation around the z axis by an angle θ measured in the direction defined by the right-hand rule, i.e., the rotation of the xy plane with the z axis fixed and the rotation angle measured in the clock-wise direction seen from the positive part of the z axis. Then, perform a rotation around the rotated y axis by an angle ϕ obeying the same rule.

In this case, the rotated coefficients

$$\frac{\Lambda_{l,m-l,n-m}(\theta,\phi)}{\sqrt{C_m^l C_n^m}} = \sum_{k=m-l}^{n}\sum_{j=0}^{k} \frac{\Lambda_{j,k-j,n-k}}{\sqrt{C_k^j C_n^k}} \cdot A_l(n-k;n-m+l;\phi)A_{m-l}(j;k;\theta)$$

satisfy $\Lambda_{n-m,m-l,l}(\theta,\phi) = \Lambda_{n-m,m}\delta_{l,0}$ if $\Lambda_{n-m,m}$ are the coefficients of a 2-D pattern moving at a speed of $u\tan\phi$ along the direction defined by θ. Furthermore, they satisfy $\Lambda_{n-m,m-l,l}(\theta,\phi) = \Lambda_n\delta_{m,0}\delta_{l,0}$ if Λ_n are the coefficients of a 1-D pattern moving at a speed of $u\tan\phi$ along its normal direction given by θ. Indeed, this rotated 3-D Hermite coefficients allows to make a classification of different patterns immersed in the spatiotemporal signal of a video sequence [43].

V. Discrete Counterpart

A. Discrete Scale-Space

For discrete 1-D signals, a complete theory can be based on a discrete analogy to the above treatment. The only nontrivial smoothing kernels of finite support arise from the generalized binomial smoothing [44], and if it is combined with a requirement that the family of smoothing transformations must obey the semi-group property $L(\cdot,\sigma)\star L(\cdot,\sigma') = L(\cdot,\sigma+\sigma')$ over scales and possesses a discrete scale parameter, then the symmetric binomial

$$B(x;N) = \frac{1}{2^{2N}}\binom{2N}{x+N} \tag{57}$$

centered at $x=0$ for $x\in\mathbb{Z}$, with the discrete scale parameter $N\in\mathbb{N}$, can be used as the discrete multiscale kernel. This kernel is termed the discrete analogue of Gaussian kernel and satisfies several properties in the discrete domain; for example, it corresponds to the discrete delta function for $N=0$, while for large N it approaches the Gaussian kernel, i.e.,

$$B(x;N) \sim G\left(x;\frac{N}{4}\right) \text{ as } N\to\infty \tag{58}$$

for $x\in\mathbb{Z}$.

B. Discrete Multiscale Derivatives

Discrete derivative approximations with scale-space properties can be built as discrete differences of the binomial kernel [45]. In this case, because the binomial kernel has a compact support, the maximum order of derivatives is limited by the scale parameter N of the binomial., i.e.,

$$\begin{aligned} B_n(x;N) &= \Delta^n B\left(x-\frac{n}{2};N-\frac{n}{2}\right) \\ &= \frac{(-1)^n}{2^{-2n}C_{2N}^n}K_n(x+N;2N)B(x;N) \end{aligned} \tag{59}$$

for $n,x=0,\ldots,N$ and $N\in\mathbb{N}$, where $K_n(x;N)$ stands for the Krawtchouk's polynomial of degree n, the discrete counterpart of Hermite polynomials [35]. These functions approach the Gaussian derivatives as N tends to infinite. Moreover, most of the properties of the Gaussian derivatives have their discrete counterpart on these binomial functions with scale parameter $N=4\sigma$. For instance, the normalized Gaussian derivatives are approximated by their discrete counterpart

$$B_n^*(x;N) = 2^{-n}\sqrt{C_N^n}B_n\left(x;\frac{N}{2}\right) \tag{60}$$

for $n=0,\ldots,N$; $x=-N/2,\ldots,N/2$ and N even. The orthogonality property (4) has its discrete form on these functions, i.e.,

$$\sum_{x=-\frac{N}{2}}^{\frac{N}{2}}\frac{B_n^*(x)B_m^*(x)}{B^*(x)} = \delta_{n,m}.$$

In order to show the most relevant properties of these discrete multiscale derivatives, we write them in the z-transform domain as

$$b_n(z;N) = z^N(1-z^{-1})^n\left(\frac{1+z^{-1}}{2}\right)^{2N-n}. \tag{61}$$

From this, their relation to the GBF introduced in previous section should be clear

$$B_n(x;N) = 2^{n-N}A_n\left(x+N;2N;\frac{\pi}{4}\right). \tag{62}$$

The *concatenation* property

$$B_n(x;N)\star B_m(x;M) = B_{n+m}(x;N+M) \tag{63}$$

is easily derived in the z-domain as the convolution operation reduces to a product operation. Also, the differentiation with respect to scale, here expressed in terms of discrete differences

$$2^{2m}\Delta_N^m\{B_n(x;N)\} = B_{2m+n}(x;N+m) \tag{64}$$

is easily verified in the z-domain. The particular case with $n=0$ and $m=1$ is recognized as the discrete counterpart of the *heat diffusion equation* on the binomial kernel [46]. If we consider the case with $n\geq 0$ and $m=1$, the following recursive form results:

$$B_n(x;N) = B_n(x;N+1) - 4^{-1}B_{n+2}(x;N+1).$$

Then, by recursive replacement of its right side for M times, the following finite expansion is obtained

$$B_n(x;N) = \sum_{m=0}^{M}(-1)^m 2^{-2m}C_M^m B_{n+2m}(x;N+M). \tag{65}$$

This finite expansion resembles the continuous Taylor expansion of a Gaussian derivative. In fact, the same limiting process that turns the discrete operators into the Gaussian derivatives turns (65) into the Taylor expansion of a Gaussian derivative.

C. Discrete Multiscale Hermite Transform

In order to arrive at a signal decomposition similar to that encountered for continuous signals, we first make the link to the continuous theory explicit. Let us define the scale sequence

$N_k = 4\sigma_k = 2^{2k+1}$ for $k = 1, 2, \ldots K$, with K integer. The scaling constant τ is maintained as in the continuous case. Then, from (65) and for $n = 0$ we can write

$$B(x; N_{k-1}) - B(x; N_k) = \sum_{n=1}^{\tau N_k} (-4)^{-n} C_{\tau N_k}^n B_{2n}(x; N_k) \tag{66}$$

which is the discrete approximation of the DoG function and it is referred to as difference of binomial function (or DoB for short). The asymptotical approximation of the DoG filters by the DoB functions holds as N_k grows indefinitely. Specifically, since

$$C_N^n = \frac{N^n}{n!}\left(1 - \frac{1}{N}\right)\cdots\left(1 - \frac{n-1}{N}\right)$$

it should be clear that

$$4^{-n} C_{\tau N_k}^n \sim \frac{\left(\frac{\tau N_k}{4}\right)^n}{n!} \text{ as } N_k \to \infty.$$

Therefore, (66) turns into (6) with $\sigma_k = N_k/4$ as N_k grows indefinitely.

As for continuous signals, we can decompose discrete signals in terms of a finite number of DoB channels; unfortunately, this approach has some limitations. The splitting of the even-order functions to produce the same filters for the analysis and synthesis stages is no longer possible. Indeed, an integer subsampling rate has to be used as the approximation for the coordinate normalization because the discrete space can not be arbitrarily scaled as the continuous space. An exact reconstruction is possible only if a weighting function is introduced in the interpolating filters as in [21]. Also, the size of the filters grows exponentially making their implementation impractical. In order to overcome these limitations, we use an alternative approach based on the pyramidal implementation of Section II-E.

By approximation of the continuous transform, the discrete counterpart of the MHT is defined in terms of the discrete filters $\Phi_n^{(1)}(x) = B_n^*(-x; 8)$ and $\Phi_n(x) = B_n^*(-x; 6)$ as follows.

First, compute the coefficients of the first layer as

$$\Lambda_n^{(1)}(q) = \sum_x L(x)\Phi_n^{(1)}(2q - x) \tag{67}$$

for $n = 0, \ldots, 8$ and $q \in \mathbb{Z}$. Then, compute the upper layers through

$$\Lambda_n^{(k+1)}(q) = \sqrt{\tau^{-n}} \sum_p \Lambda_0^{(k)}(p)\Phi_n(2q - p) \tag{68}$$

for $k = 0, \ldots, K-1$ and $n = 0, \ldots, 6$. Since these discrete coefficients approach the continuous ones within an order of approximation given by the resolution level, one may insert these coefficients into any of the expressions derived for the continuous coefficients and expect an approximate behavior. For instance, the second-order gauge condition given by (45) was combined with (39), (55), and (27) to compute different multiscale decompositions of the well-known HOUSE image. The original image and the different representations are displayed in Fig. 6 for illustration purposes. Notice that the highly oriented patterns dominating the image lead to a good energy compactness along the rotated coordinate x as predicted for continuous signals [Fig. 6(c)]. In this case, the coefficient $\Lambda_{1,1} = 0$ for every resolution level. Although no significant differences are observed in the expansions of Fig. 6(c) and (d), it should be noted that they are indeed the same for the first order and differ more as the order increases.

Analogous to the continuous case, the inverse transform is defined by the following interpolation process:

$$\Lambda^{(k)}(q) = \sum_{n=0}^{6} \sqrt{\tau^n} \sum_p \Lambda_n^{(k+1)}(p) 2\Phi_n(2p - q) \tag{69}$$

for $k = K-1, \ldots, 1$ and

$$L(x) = \sum_{n=0}^{8} \sum_q \Lambda_n^{(1)}(q) 2\Phi_n^{(1)}(2q - x) \tag{70}$$

for the signal reconstruction. The reconstruction is perfect provided that

$$\sum_{n=0}^{N} \sum_p 2B_n^*(q - 2p; N) B_n^*(r - 2p; N) = \delta_{q,r}$$

for $p, q, r \in \mathbb{Z}$ and $N > 0$.

D. Practical Issues and Applications

So far, we have focused mainly on the theoretical derivations of the transform and its properties. Providing an extensive description of every application of the model presented here is beyond the scope of this article. We rather discuss some practical issues and suggest how this transform might be applied to digital image processing.

A very natural question to ask is what is the effect of limiting the maximum order of derivation in the signal reconstruction? Or, how significant are the errors introduced due to such approximation? Of course, the answer will depend on the kind of signal and its purpose. For image representation, we can use the pseudo signal-to-noise ratio (PSNR) as a measure of the approximation. The PSNR of several reconstructed images are plotted against the maximum order in Fig. 7. The exact reconstruction can be achieved only when all the coefficients are used, e.g., 16 in the 2-D space. A *reasonable* PSNR (say above 30 dB) is most likely to be achieved for a maximum order of 8. This number can be lower for the upper layers and even lower if based on purely subjective criteria since the blurred effect introduced by this transform is more tolerable by human viewers than other effects, e.g., the ringing produced by a wavelet decomposition.

As we noticed for the continuous case, the MHT representation requires the low-pass ($n = 0$) coefficient only at layer $k = K$, for the signal reconstruction. Fortunately, at this resolution level, the signal rate is reduced by a factor of 2^K with respect to the input rate. In general, a single coefficient at level k has a size reduced by a factor of 2^k with respect to the input signal so that the whole representation will have a size increased by a factor of $7 - 5 \times 2^{-K}$ with respect to the input signal. This factor reaches its minimum value for the case $K = 1$ (single-scale case) and for larger values of K, it approaches 7. This redundancy could be a major inconvenience for some applications such as coding. In that case, one can predict approximately the coefficients from the coarser resolution level based

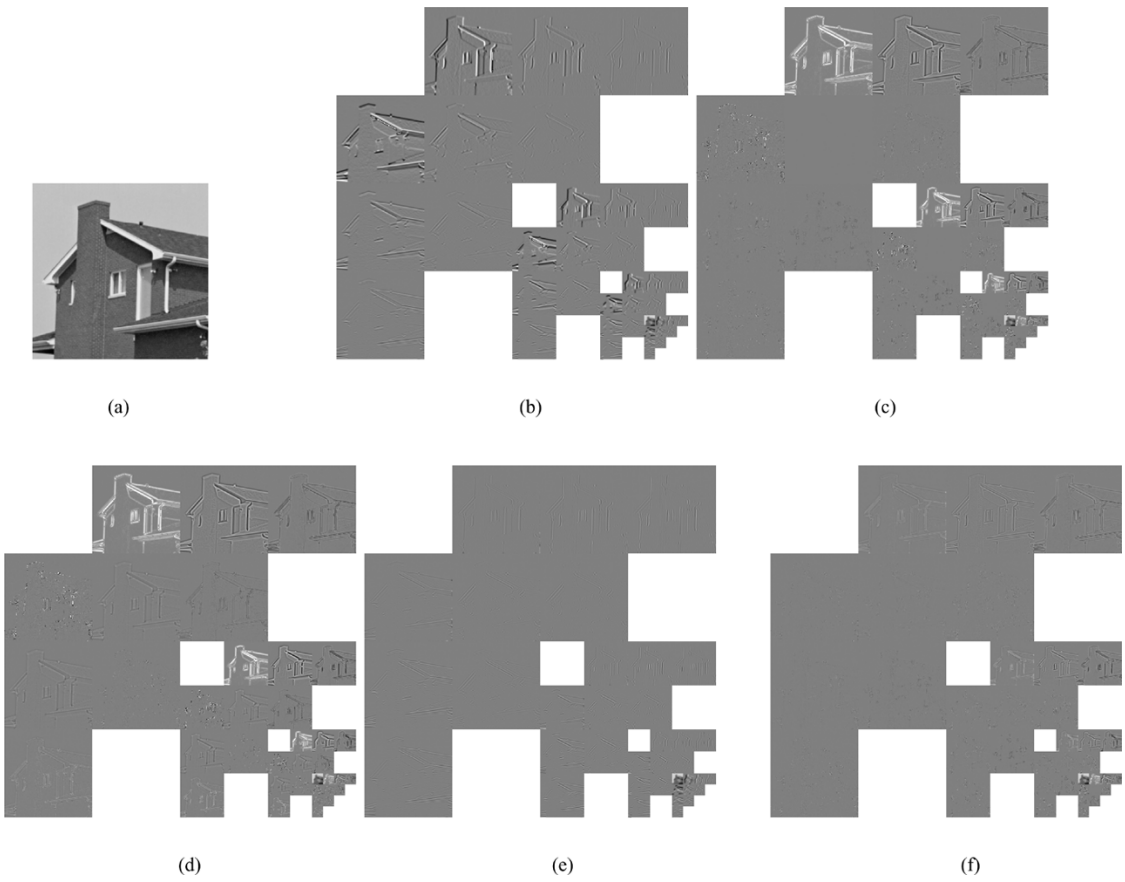

Fig. 6. (a) HOUSE image. (b)–(f) Discrete multiscale decompositions up to third order and $K = 4$. The coefficients are arranged in increasing order from left to right in x and from top to bottom in y. The orientations used in (c), (d), and (f) were determined from the second-order gauge condition (47). The full expansion was used in (e) and (f) for the prediction of coarser levels (see the text). (a) *input*; (b) *original coefficients*; (c) *rotated coefficients*; (d) *multidirectional*; (e) *prediction errors*; and (f) *rotated errors*.

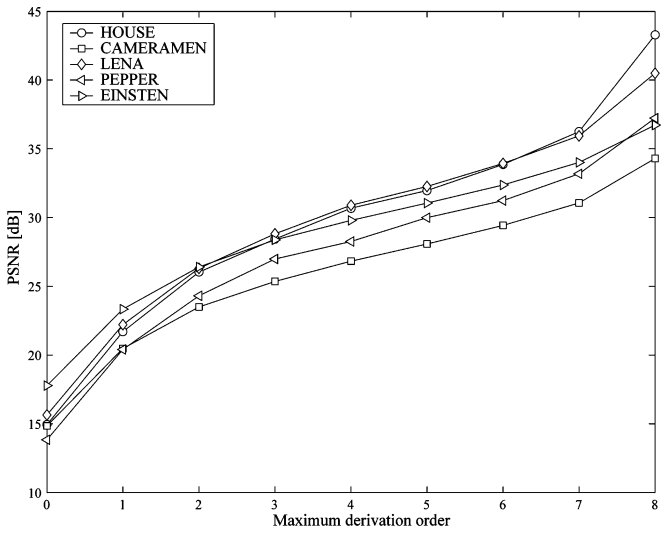

Fig. 7. Effect of limiting the maximum derivation order of the 2-D expansion (with $K = 4$) on the PSNR of several reconstructed images.

on (27); then, the residual can be further rotated and encoded accordingly. Fig. 6(e) and (f) illustrates the residual errors obtained for the image HOUSE without and with local rotation, respectively. In both cases, prediction errors tend to be higher for higher order coefficients; however, the rotated version tends to compact the errors along the first dimension.

In Section III-B, we introduced a method to perform the anisotropic Gaussian diffusion of images. It requires estimating the orientation and scales along each orthogonal direction defined by the orientation angle at every position in the image. To illustrate how it works, we consider a single scale decomposition and classify the patterns at each sampling location as 0-D, 1-D, and 2-D patterns as in [29]. The scales for 0-D patterns should equal scale of the representation (therefore, $\tau_u = \tau_v = 0$) and the angle can be arbitrarily selected as it does not play a role. For 1-D patterns, the orientation is taken as the gradient angle. The scale along the gradient direction is preserved as in the input signal $(\tau_u = 1)$, while the scale along the orthogonal direction is taken as the scale of the representation $(\tau_v = 0)$. For 2-D patterns, both scales are preserved $(\tau_u = \tau_v = 1)$ and the orientation can be arbitrarily selected. Fig. 8 shows the results obtained with several images. A similar setting can be applied to image coding (cf. [47]).

VI. Conclusion

We have introduced the multiscale hermite transform and presented new theoretical results on local orientation analysis based on this transform. The major properties of the signal decomposition were studied in 1-D for simplicity, but they are straightforwardly generalized to a multidimensional separable transform. Moreover, since orientation analysis makes sense only for data

Fig. 8. (a) Input images, (b) enlarged portion of original images, (c) enlarged portion of diffused images, and (d) plots of (dots) 0-D patterns, (line segments) 1-D patterns, and (circles) 2-D patterns. A single-scale decomposition was used. The angle was set to the gradient orientation. The smoothing parameters (γ_u, γ_v) were set to (0,0), (1,0) and (1,1) for 0-D, 1-D, and 2-D patterns, respectively (see the text).

in two and higher dimensions, the major results in this matter where performed in 2-D and indications for its application to 3-D data sets, such as video sequences, were provided.

The basic operators of this representation are Gaussian derivatives, which have been largely studied in the literature both from the point of view of their utility in computer vision [18], [48], [49] and for their relevance in the modeling of the human visual system [10], [12], [13], [50], [51]. The Hermite transform was originally introduced as a broader theoretical framework for these approaches [21], [23], [25]. In our case, the multiscale decomposition resulted from the explicit choice of certain channels resembling the ones found in the human visual system, namely, the difference of Gaussians [26], [27]. These channels were further decomposed into Gaussian derivatives which are also relevant in the modeling of the human visual system and have many symmetry properties facilitating the multiscale and multiorientation analyses.

The major contribution of this paper relative to previous work is two-fold. First, we took advantage of the multiscale properties of the basis functions to build an efficient implementation through a pyramidal approach. In the continuous setting, the transform is steerable in both scale and orientation domains, i.e., one can reconstruct the coefficients at any orientation and scale from discrete samples of these domains [6], [7]. As an alternative, one can predict finer levels from coarser levels of the pyramid. The other contribution is respect to the implementation of the steering coefficients required to rotate the Gaussian derivatives. We identified these coefficients as a generalized binomial family and derived several properties on them. The rotation is then performed efficiently by means of these discrete sequences in a cascaded fashion. The application to local orientation analysis was also discussed. We demonstrated how the Cartesian coefficients can be mapped from or to multidirectional coefficients.

The discrete MHT exhibits many advantages for adaptive image processing.

1) It is efficiently computed. The binomial filters can be performed with a time complexity in the order of $(N+1)\log_2(N+1)$ per sample [41].
2) It approximates the behavior of the continuous transform. The binomial filters are a complete family perse [16] with the unique property of being linked to the Gaussian family through the scale parameter. The degree of approximation is directly related to this parameter.
3) It allows to approximate the rotated MHT in a very efficient way. The GBF defines a self-inverting discrete transform on the partial derivation order that can be applied to the discrete coefficients without explicitly rotating the binomial filters. A perfect reconstruction of the original coefficients is possible. The multidirectional case is not perfectly invertible, however, the reconstruction errors may be neglected for some applications.
4) It can be inserted in a predictive scheme. The higher resolution coefficients, particularly the ones of low order (e.g., $n = 1,2,3,4$), can be approximately recovered from the next coarser resolution level. This may be useful for coding applications.
5) It can be applied to broad variety of image processing. For instance, it can be efficiently applied to optic flow estimation, where filtering plays an important role [52].

Spatiotemporal filtering can be particularly time- and memory consuming if done with inseparable filters. We have demonstrated how directional derivatives along rotated 3-D coordinates can be efficiently computed through the GBF.

6) It may be linked to other efficient representations such as quadrature mirror filters and wavelets [53].

APPENDIX I
Z TRANSFORM OF THE GBF

After developing the binomials in (36) we can write

$$a_n(z;N;\theta)=\sum_{k=0}^{N-n}\sum_{i=0}^{n}(-1)^i C_n^i C_{N-n}^k c^{n-i+k}s^{N-n+i-k}z^{-i-k}$$

with $s=\sin\theta$ and $c=\cos\theta$. Then, replacing $k+i$ by k while fixing the sums limits we have

$$\begin{aligned}
&a_n(z;N;\theta)\\
&\quad=\sum_{k=0}^{N}z^{-k}\sum_{i=0}^{n}(-1)^i C_n^i C_{N-n}^{k-i}c^{n-2i+k}s^{N-n+2i-k}\\
&\quad=\sum_{k=0}^{N}z^{-k}\sum_{j=0}^{n}(-k)^{n-j}C_n^j C_{N-n}^{k+j-n}c^{2j+k-n}s^{N+n-2j-k}\\
&\quad=\sum_{k=0}^{N}z^{-k}\left(\frac{s}{c}\right)^k\sum_{j=0}^{n}(-1)^{n-j}C_n^j C_{N-n}^{k+j-n}c^{2(k+j)-n}\\
&\qquad\cdot s^{N+n-2(k+j)}\\
&\quad=\sum_{k=0}^{N}z^{-k}\left(\frac{s}{c}\right)^k\Delta^n\left\{C_{N-n}^{k-n}c^{2k-n}s^{N+n-2k}\right\}
\end{aligned}$$

where the nth forward difference operator Δ^n is applied over the index k. From the z-transform definition, it is clear that the sequence after z^{-k} corresponds to the impulse response of the GBF.

APPENDIX II
GBF FOR A DIFFERENCE ANGLE

We use the notation s and c for the sine and cosine functions respectively and denote the argument by a subscript. Then, after inserting the sine and cosine of the difference of two angles into (36), we can write

$$\begin{aligned}
a_n(z;N;\theta-\phi)&=\left(c_\theta c_\phi+s_\theta s_\phi-z^{-1}(s_\theta c_\phi-c_\theta s_\phi)\right)^n\\
&\quad\cdot\left(s_\theta c_\phi-c_\theta s_\phi+z^{-1}(c_\theta c_\phi+s_\theta s_\phi)\right)^{N-n}\\
&=\left(c_\phi(c_\theta-z^{-1}s_\theta)+s_\phi(s_\theta+z^{-1}c_\theta)\right)^n\\
&\quad\cdot\left(c_\phi(s_\theta+z^{-1}c_\theta)-s_\phi(c_\theta-z^{-1}s_\theta)\right)^{N-n}\\
&=(s_\theta+z^{-1}c_\theta)^N\left(c_\phi-s_\phi\frac{c_\theta-z^{-1}s_\theta}{s_\theta+z^{-1}c_\theta}\right)^{N-n}\\
&\quad\cdot\left(s_\phi+c_\phi\frac{c_\theta-z^{-1}s_\theta}{s_\theta+z^{-1}c_\theta}\right)^n
\end{aligned}$$

where the last binomials in the right term are similar to those appearing in the GBF transfer function and therefore they must develop into

$$\sum_{m=0}^{N}\left(\frac{c_\theta-z^{-1}s_\theta}{s_\theta+z^{-1}c_\theta}\right)^m A_{N-n}(m;N;\phi)$$

according to Appendix I. Therefore, we can write

$$\begin{aligned}
a_n(z;N;\theta-\phi)&=\sum_{m=0}^{N}(c_\theta-z^{-1}s_\theta)^m(s_\theta+z^{-1}c_\theta)^{N-m}\\
&\quad\cdot A_{N-n}(m;N;\phi)\\
&=\sum_{k=0}^{N}z^{-k}\sum_{m=0}^{N}A_m(k;N;\theta)A_{N-n}(m;N;\phi)
\end{aligned}$$

where, once again, the binomials are expanded into the transfer function of the GBF with the appropriate index. Therefore

$$A_n(k;N;\theta-\phi)=\sum_{m=0}^{N}A_m(k;N;\theta)A_{N-n}(m;N;\phi)\quad(71)$$

follows from the z transform definition. Furthermore, if we set $\theta=\phi$, the orthogonality property

$$\sum_{m=0}^{N}A_m(k;N;\theta)A_n(m;N;\theta)=\delta_{k,n}\quad(72)$$

results.

APPENDIX III
RELATION OF GBF TO TRIGONOMETRIC FUNCTIONS

The angular function defined by the GBF are readily related to trigonometric functions. First, from the Euler identity and the binomial expansion, we can write

$$\begin{aligned}
j^n e^{-jN\theta}&=e^{-j\left(N\theta-\frac{\pi}{2}n\right)}\\
&=e^{-j\theta(N-n)}e^{-j\left(\theta-\frac{\pi}{2}\right)n}\\
&=(\cos\theta-j\sin\theta)^{N-n}(\sin\theta+j\cos\theta)^n\\
&=\sum_{k=0}^{N}j^k A_{N-n}(k;N;\theta)\\
\Rightarrow e^{-jN\theta}&=\sum_{k=0}^{N}j^{k-n}A_{N-n}(k;N;\theta)\\
&=\sum_{k=0}^{N}j^{n-k}A_n(N-k;N;\theta)
\end{aligned}$$

where we have used the property (37a) in this last equality. Then, by separation of the real and complex parts results in

$$\cos(N\theta)=\sum_{k=0}^{\lfloor\frac{N}{2}\rfloor}(-1)^{n-k}A_{2n}(N-2k;N;\theta)\quad(73a)$$

for $n = 0, \ldots, \lfloor N/2 \rfloor$, or

$$\cos(N\theta) = \sum_{k=0}^{\lfloor \frac{N-1}{2} \rfloor} (-1)^{n-k} A_{2n+1}(N-2k-1; N; \theta) \quad (73b)$$

for $n = 0, \ldots, \lfloor (N-1)/2 \rfloor$, from the real part, and

$$\sin(N\theta) = \sum_{k=0}^{\lfloor \frac{N-1}{2} \rfloor} (-1)^{n-k} A_{2n}(N-2k-1; N; \theta) \quad (74a)$$

for $n = 0, \ldots, \lfloor N/2 \rfloor$, or

$$\sin(N\theta) = -\sum_{k=0}^{\lfloor \frac{N}{2} \rfloor} (-1)^{n-k} A_{2n+1}(N-2k; N; \theta) \quad (74b)$$

for $n = 0, \ldots, \lfloor (N-1)/2 \rfloor$, from the complex part.

On the other hand, if we replace the identities

$$\cos\theta = \frac{e^{j\theta} + e^{-j\theta}}{2} \quad \text{and} \quad \sin\theta = \frac{e^{j\theta} - e^{-j\theta}}{2j}$$

into (36), we can simplify and take the inverse z-transform to find

$$A_n(k; N; \theta) = \sum_{m=0}^{N} j^{N-2m-n-k} A_n(m; N) A_m(k; N) e^{j\theta(2m-N)}$$

where $A_n(k; N) = A_n(k; N; \pi/4)$ defines the symmetric binomial family (SBF). Furthermore, since we can set $A_n(k; N) = 0$ for non integer values of n and k the change of index from $2m - N$ to m is feasible. Then, we can write this expansion in the form of the finite Fourier series

$$A_n(k; N; \theta) = \sum_{m=-N}^{N} c_m e^{jm\theta} \quad (75a)$$

with the Fourier coefficients given by

$$c_m = \frac{1}{2\pi} \int_0^{2\pi} A_n(k; N; \theta) e^{-jm\theta} d\theta = j^{-m-n-k} A_n\left(\frac{N}{2} + \frac{m}{2}; N\right) A_{\frac{N}{2}+\frac{m}{2}}(k; N) \quad (75b)$$

for $m = -N, \ldots, N$.

ACKNOWLEDGMENT

The authors would like to thank their respective institutions for the support provided and to the anonymous reviewers for their valuable comments and corrections.

REFERENCES

[1] S. Marcelja, "Mathematical description of the responses of simple cortical cells," *J. Opt. Soc. Amer.*, vol. 70, pp. 1297–1300, 1980.
[2] J. G. Daugman, "Two-dimensional spectral analysis of cortical receptive fields profiles," *Vis. Res.*, vol. 20, pp. 847–856, 1980.
[3] M. Bastiaans, "Gabor's signal expansion and degrees of freedom of a signal," *Opt. Acta*, vol. 29, pp. 1223–1229, 1982.
[4] M. Porat and M. Zeevi, "The generalized gabor scheme of image representation in biological and machine vision," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 10, no. 4, pp. 452–467, Jul. 1988.
[5] A. B. Watson, "The cortex transform: rapid computation of simulated neural images," *Comput. Vis., Graph., Image Process.*, vol. 39, pp. 311–327, 1987.
[6] W. T. Freeman and E. H. Adelson, "The design and use of steerable filters," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 13, no. 9, pp. 891–906, Sep. 1991.
[7] E. P. Simoncelli, W. T. Freeman, E. H. Adelson, and D. J. Heeger, "Shiftable multiscale transforms," *IEEE Trans. Inf. Theory*, vol. 38, no. 2, pp. 587–607, Mar. 1992.
[8] E. P. Simoncelli and W. T. Freeman, "The steerable pyramind: A flexible architecture for multi-scale derivative computation," in *Proc. Int. Conf. Image Process.*, 1995, pp. 444–447.
[9] G. H. Granlund and H. Knutsson, *Signal Processing for Computer Vision*. Dordrecht, The Netherlands: Kluwer, 1995.
[10] R. A. Young, "The Gaussian Derivative Theory of Spatial Vision: Analysis of Cortical Receptive Field Line-Weighting Profiles," Gen. Motors Res., Tech. Rep. GRM-4920, 1985.
[11] ——, "Simulation of human retinal function with the gaussian derivative model," in *Proc. IEEE CCVP*, Miami, FL, 1986, pp. 564–569.
[12] ——, "The gaussian derivative model for spatial vision: I. retinal mechanisms," *Spat. Vis.*, vol. 2, pp. 273–293, 1987.
[13] J. J. Koenderink and A. J. V. Doorn, "Representation of local geometry in the visual system," *Biol. Cybern.*, vol. 55, pp. 367–375, 1987.
[14] ——, "Generic neighborhood operators," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 14, no. 6, pp. 597–605, Jun. 1992.
[15] A. Witkin, "Scale-space filtering: A new approach to multiscale description," *Image Understand.*, vol. 3, pp. 79–95, 1984.
[16] R. A. Haddad and A. N. Akansu, "A new orthogonal transform for signal coding," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 36, no. 9, pp. 1404–1411, Sep. 1988.
[17] A. P. Morgan, L. T. Watson, and R. A. Young, "A Gaussian derivative based version of JPEG for image compression and decompression," *IEEE Trans. Image Process.*, vol. 7, no. 9, pp. 1311–1320, Sep. 1998.
[18] T. Lindeberg, "Edge detection and ridge detection with automatic scale selection," *Int. J. Comput. Vis.*, vol. 30, no. 2, pp. 117–156, 1998.
[19] M. Jacob and M. Unser, "Design of steerable filters for feature detection using canny-like criteria," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 26, no. 8, pp. 1007–1019, Aug. 2004.
[20] R. A. Haddad, "A class of orthogonal nonrecursive binomial filters," *IEEE Trans. Audio Electroacoust.*, vol. AV-19, no. 4, pp. 296–304, Dec. 1971.
[21] J. B. Martens, "The hermite transform-theory," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 38, no. 9, pp. 1595–1606, Sep. 1990.
[22] J. A. Bloom and T. R. Reed, "A Gaussian derivative-based transform," *IEEE Trans. Image Process.*, vol. 5, no. 3, pp. 551–553, Mar. 1996.
[23] J. B. Martens, "The Hermite transform-applications," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 38, no. 9, pp. 1607–1618, Sep. 1990.
[24] M. Michaelis and G. Sommer, "A lie group-approach to steerable filters," *Pattern Recognit. Lett.*, vol. 16, no. 11, pp. 1165–1174, 1995.
[25] J. B. Martens, "Local orientation analysis in images by means of the Hermite transform," *IEEE Trans. Image Process.*, vol. 6, no. 8, pp. 1103–1116, Aug. 1997.
[26] R. W. Rodieck, "Quantitative analysis of cat retinal ganglion cell response to visual stimuli," *Vis. Res.*, vol. 5, pp. 583–601, 1965.
[27] R. V. Rullen and S. J. Thorpe, "Rate coding versus temporal order coding: what the retinal ganglion cells tell the visual cortex," *Neural Comput.*, vol. 13, no. 6, pp. 1255–1283, 2001.
[28] J. L. Silván-Cárdenas and B. Escalante-Ramírez, "A multi-scale, multi-orientation image transform," in *Proc. 5th World Multiconference on Systemics, Cybernetics, and Informatics*, Orlando, FL, 2001, pp. 574–578.
[29] ——, "Image coding with directional-oriented hermite transform on a hexagonal lattice," in *Proc. Applications of Digital Image Processing XXIV*, vol. 4472, A. G. Tescher, Ed., San Diego, CA, 2001, pp. 528–536.

[30] B. Escalante-Ramírez and J. L. Silván-Cárdenas, "A multiresolution directional-oriented image transform based on gaussian derivatives," in *Proc. Wavelets: Applications in Signal and Image Processing IX*, vol. 4478, M. A. U. A. F. Laine and A. Aldroubi, Eds., San Diego, CA, 2001, pp. 315–322.
[31] P. J. Burt and E. H. Adelson, "The Laplacian pyramid as a compact image code," *IEEE Trans. Commun.*, vol. COM-31, no. 4, pp. 532–540, Apr. 1983.
[32] J. J. Koenderink, "The structure of images," *Biol. Cybern.*, vol. 50, pp. 363–370, 1984.
[33] W. R. Hendee and P. N. T. Wells, *The Perception of Visual Information*. New York: Springer-Verlag, 1993, ch. 4, pp. 96–97.
[34] M. Felsberg and G. Sommer, "Structure multivector for local analysis of images," in *Multi-Image Analysis*, R. Klette, T. Huang, and G. Gimel'farb, Eds. Berlin, Germany: Springer-Verlag, 2001, vol. 2032, pp. 93–104.
[35] G. Szegö, *Orthogonal Polynomials*: Amer. Math. Soc., 1959.
[36] B. Escalante-Ramírez and J. B. Martens, "Noise reduction in computerized tomography images by means of polynomial transforms," *Vis. Commun. Image Represent.*, vol. 3, pp. 272–285, 1992.
[37] I. Daubechies, "Orthonormal bases of compactly supported wavelets," *Commun. Pure Appl. Math.*, vol. 41, pp. 909–996, 1988.
[38] S. G. Mallat, "Multifrequency channel decompositions of images and wavelet models," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 37, no. 12, pp. 2091–2110, Dec. 1989.
[39] A. A. Bharath, "Steerable filters from erlang functions," in *Proc. Brit. Machine Vision Conf.*, vol. 1, 1998, pp. 144–153.
[40] M. Felsberg and G. Sommer, "The monogenic scale-space: a unifying approach to phase-based image processing in scale-space," *J. Math. Imag. Vision*, vol. 21, pp. 5–26, 2004.
[41] M. Hashimoto and J. Sklansky, "Multiple-order derivatives for detecting local image characteristics," *Comput. Vis., Graph., Image Process.*, vol. 39, pp. 28–55, 1987.
[42] P. Camarillo-Sandoval, A. Varela-López, and B. Escalante-Ramírez, "Adaptive multiplicative-noise reduction in SAR images with polynomial transforms," in *Proc. IGARSS*, Seattle, WA, 1998, pp. 1171–1173.
[43] B. Escalante-Ramírez and J. L. Silván-Cárdenas, "Motion analysis and classification with directional Gaussian derivatives in image sequences," in *Proc. Advanced Signal Processing Algorithms, Architectures and Implementations*, vol. 4116, F. T. Luk, Ed., San Diego, CA, 2000, pp. 447–453.
[44] T. Lindeberg, "Scale-space for discrete signals," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 12, no. 3, pp. 234–254, Mar. 1990.
[45] ——, "Discrete derivative approximations with scale-space properties: A basis for low-level feature extraction," *J. Math. Imag. Vis.*, vol. 3, pp. 349–376, 1993.
[46] R. Hummel and R. Moniot, "Reconstruction from zero crossings in scale space," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 37, no. 12, pp. 2111–2130, Dec. 1989.
[47] A. M. V. Dijk and J. B. Martens, "Image representation and compression with steered hermite transform," *Signal Process.*, vol. 56, no. 1, pp. 1–16, 1997.
[48] R. Hartley, "A Gaussian-weighted multiresolution edge detector," *Comput. Vis., Graph., Image Process.*, vol. 30, pp. 70–83, 1985.
[49] D. Ziou, S. Wang, and J. Villancourt, "Depth from defocus using the Hermite transform," in *Proc. Int. Conf. Image Processing*, Chicago, IL, 1998, pp. 958–962.
[50] D. Marr, *Vision: A Computational Investigation Into the Human Representation and Processing of Visual Information*. New York: Freeman, 1982.
[51] R. E. Frye and R. S. Ledley, "Derivative of gaussian functions as receptive field models for disparity sensitive neurons of the visual cortex," in *Proc. 15th Southern Biomedical Engineering Conf.*, 1996, pp. 270–273.
[52] D. J. Fleet and A. D. Jepson, "Hierarchical construction of orientation and velocity selective filters," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 11, no. 3, pp. 315–325, Mar. 1989.
[53] A. N. Akansu, R. A. Haddad, and H. Caglar, "The binomial QMF-wavelet transform for multiresolution signal decomposition," *IEEE Trans. Signal Process.*, vol. 41, no. 1, pp. 13–19, Jan. 1993.

José L. Silván-Cárdenas was born in Campeche, Mexico, in 1973. He received the B.S. (Hon.) degree in computer engineering and the M.S. (Hon.) degree in electrical engineering degree from the National University of Mexico (UNAM), in 1998 and 2002, respectively. His main research interest at the time was image coding and image analysis based on models of the human visual system.

Since July 2001, he has been with the Geography and Geomatic Research Center (CIGGET), Tlalpan, Mexico. His current research interests include image classification for land cover mapping, automatic road detection from multispectral imagery, and artificial intelligence techniques applied to digital image processing.

Boris Escalante-Ramírez was born in Mexico City in 1961. He received the B.S. degree in electrical engineering from the National University of Mexico (UNAM) in 1985, the M.S. degree in electronic engineering degree from the Philips International Institute of Technological Studies, Eindhoven, The Netherlands, in 1987, and the Ph.D. degree from the Eindhoven University of Technology in 1992.

Since then, he has been with the Graduate Division of the School of Engineering, UNAM. His research interests are visual information processing, perception models and their applications to image processing including restoration, coding, segmentation, fusion, and optic flow.