

The Hermite Transform—Theory

JEAN-BERNARD MARTENS

Abstract—A digital image is usually specified by an array of point-wise intensities. For an intelligent interpretation of the image data we need to make the important information explicit. This usually implies determining spatiotemporal relationships between these intensities, and hence requires some form of local processing of the visual data. In this paper we introduce a new scheme for the local processing of visual information, called the Hermite transform. We have addressed the problem from the point of view of image coding, and therefore the proposed scheme is presented as an analysis/resynthesis system. Our objectives are, however, not restricted to coding. First, the analysis part is designed so that it can also serve applications in the area of computer vision. Indeed, derivatives of Gaussians, which have found widespread application in feature detection over the past few years, play a central role in the Hermite analysis. Second, next to integrating ideas from the distinct, but related, areas of image coding and computer vision, it is argued that the proposed processing scheme is also in close agreement with our current insight into the image processing that is carried out by the human visual system.

I. INTRODUCTION

FOR many applications in image coding and computer vision, as well as in the case of human visual perception, it is required that the image data, which are given as an array of intensity values, be interpreted into meaningful visual patterns. It is generally agreed that some form of local spatiotemporal processing of the original data is required for that purpose. This kind of processing involves two important decisions. First, to make the processing local, the image is usually multiplied by a window function. The size of the window establishes the set of image points that contribute to one basic processing step. The form of the window function determines the relative weight of each contributing image point. In order to describe the image completely, this local processing has to be repeated for a sufficient number of window positions. The form, size, and spacing of the window function have to be selected. Second, for each position of the window, specific processing steps have to be undertaken. As any specific choice of processing implies the search for specific patterns, selecting this process is equivalent to fixing the visual patterns which are considered most relevant *a priori*.

It is very difficult to make optimal choices for the window function and processing on purely theoretical arguments. The human visual system is therefore often used as a reference, an approach that is also adopted in this paper. It is, however, interesting to briefly review avail-

able image processing techniques with respect to the choices that have been made for the window function and processing.

A class of window functions that are often used are square, nonoverlapping windows. They are, for instance, adopted in transform coding (TC) [1] and image vector quantization (VQ) [2], and represent the simplest way of subdividing an image. An important disadvantage is that at low data rates, blocking effects appear which are quite annoying to a human observer. Difference pulse code modulation (DPCM) [3], on the other hand, uses some form of scanning pattern together with a prediction window that only includes preceding pixels. Hence, only part of the neighboring information is included in the image analysis. Both preceding window functions represent fairly artificial subdivisions of the pixel domain, and have little perceptual relevance. Most second generation image coding techniques [4] such as pyramid coders (PC) [5], subband coders (SBC) [6]–[8], the cortex transform [9], [10], and Gabor expansions [11], [12], use overlapping window functions. Recent papers show that several of these coding techniques (PC, SBC), which are most often described in the frequency domain, can also be described in the spatial domain by wavelet theory [13], [14]. Wavelet theory shows how signals can be expanded on a family of functions which are the dilate and translate of a unique (window) function. The window functions that satisfy all necessary orthogonality conditions do, however, have two disadvantages. First, they are usually much larger than the window spacing (a window of length 32 for a subsampling factor of 2 is not uncommon in SBC [8]). Second, they often have considerable ringing and are hence less smooth than the window functions deduced from overlapping receptive fields in human visual perception [15]. In order to get smooth window functions, it is necessary to drop the orthogonality condition between basis functions of adjoining windows. This is, for instance, done in Gabor expansions, where smooth (Gaussian) window functions are used. There is, however, no need for dropping the orthogonality condition between the basis functions belonging to one window. The alternative image description method presented in this paper differs from Gabor expansions in this aspect. The effect of using (smooth) overlapping window functions is that images become unsharp at low data rates. Human observers usually consider this kind of image degradation more natural and less annoying than blocking effects [16]. The comparison is very similar to nearest neighbor versus higher order interpolation [17].

Manuscript received May 10, 1988; revised October 5, 1989.

The author is with the Institute for Perception Research, 5600 MB Eindhoven, The Netherlands.

IEEE Log Number 9036882.

An important parameter of a window function is its size (or scale). The selection of an appropriate window size poses a fundamental problem. On the one hand, in order to enable high data reductions, the window size has to be sufficiently large. On the other hand, the complexity of the analysis within each window increases rapidly with the window size. There are two possible approaches to the problem. First, we can select a window of fixed size and perform an analysis within each window that is sufficiently complex to include all visual patterns of interest. This approach is taken in many available coding techniques. Second, we can limit the complexity of the analysis that we perform in each window, and subsequently determine the window size needed to describe the image locally with sufficient accuracy. Hence, instead of restricting the processing to one scale, we repeat the same processing at multiple scales and subsequently use the outputs of this processing stage to select the optimum scale at each position [18]–[21]. There is compelling evidence that the human visual system uses this “scale space” principle [22], and therefore we will also adopt it in our approach.

If window functions of different sizes are used, then the spacing of the window functions is usually taken proportional to their size. In most pyramid structures [5], [23], the window spacing and size increase in steps of two. It has been shown that this exponential increase in window size and spacing is not only preferable from an information-theoretical point of view [24], [25], but is also computationally efficient [26], [27].

Last but not least, we briefly review different local signal processing techniques that have been developed in the past. Statistical coding methods such as DPCM and VQ do not give an *a priori* specification of the basic visual patterns that are used in the analysis. Instead, they use the statistics of “typical” images to determine these patterns by iterative techniques. Although these methods were first applied directly to the image data, nowadays they are usually combined with other preprocessing techniques. This acknowledges the fact that describing the images by *a priori* selected patterns, and subsequently using statistical methods to model the remaining dependencies, is often more fruitful than a direct statistical approach [8]. It also limits the complexity of the statistical search. In this paper, we concentrate on the preprocessing stage. Wavelet decompositions (with PC and SBC as special cases) fit a selected spatial pattern to the image. However, as stated earlier, the patterns that are typically used do not correlate very well with the filtering that is found in human vision. Using patterns with different orientations, as is done in directional coding [28]–[30] and the cortex transform, is in close agreement with human visual perception, and will also be included in our approach. In TC and Gabor expansions, a local harmonic analysis is performed on the signal. These harmonic descriptions are used extensively in image coding, but have found little application in image analysis. In computer vision, the image processing problem is addressed from the point of view of interpre-

tation, as one is mainly interested in finding the important image features such as edges and lines [31]–[33]. The resulting processing often involves the use of first- and second-order derivatives, almost always in combination with some regularizing low-pass filter [34]. In this paper, we introduce a new signal transformation technique that involves these operators. One advantage is that the interpretation of this new transform can profit from the vast experience with such operators in computer vision. Moreover, it was demonstrated recently [35]–[37] that derivatives of Gaussians can model filter operations in human vision with the same accuracy as the often used Gabor filters. The derivatives of Gaussian operators even have the advantage that they accomplish this modeling with fewer parameters. Hence, the new transform is also supported by current insights into human visual perception.

In Section II, we define a broad class of signal representations based on polynomial approximations within a local window. These representations will be called polynomial transforms. In Section III, we argue the importance of Gaussian windows. The resulting transform for this specific choice of window is called the Hermite transform. Its most interesting properties are also derived. The extension of the one-dimensional (1D) theory of Section II to two dimensions is described in Section IV. It is shown how directional selectivity can be introduced in two-dimensional (2D) polynomial transforms. In Section V we discuss some of the simplifications that arise in the case of the 2D Hermite transform. The three-dimensional case is discussed shortly in Section VI, where we concentrate mainly on the velocity selectivity property of the transform. In Section VII, we show how polynomial transforms can be reformulated for discrete signals. The discrete Hermite transform is defined in Section VIII to correspond to binomial windows.

In the accompanying paper [51], we will discuss the performance of the Hermite transform in a multiscale analysis. Applications in image coding and computer vision will also be illustrated in that paper.

II. ONE-DIMENSIONAL POLYNOMIAL TRANSFORMS

In this section we develop a new signal decomposition technique, called a polynomial transform, in which signals are locally approximated by polynomials. We introduce the basic ideas on 1D analog signals. Extensions to multiple dimensions and discrete signals will be considered in the following sections.

The analysis by a polynomial transform involves two steps. In a first step the original signal $L(x)$ is localized by multiplying it by a window function $V(x)$. A complete description of the signal requires that the localization process is repeated at a sufficient number of window positions. We consider the case of equidistant spacing between windows.

From the localized window function $V(x)$, we can construct a *weighting* function

$$W(x) = \sum_k V(x - kT) \quad (1)$$

by periodic repetition. The weighting function is itself periodic with period T . Provided $W(x)$ is nonzero for all x , we get

$$L(x) = \frac{1}{W(x)} \sum_k L(x) \cdot V(x - kT) \quad (2)$$

so that we are guaranteed that the localized signals $L(x) \cdot V(x - kT)$ for all different window positions kT contain sufficient information about the original signal.

The second step consists of approximating the signal piece within the window $V(x - kT)$ by a polynomial. As basis functions for the polynomial expansion, we take the polynomials $G_n(x)$, degree $[G_n(x)] = n$, that are orthonormal with respect to $V^2(x)$, i.e.,

$$\int_{-\infty}^{+\infty} V^2(x) G_m(x) G_n(x) dx = \delta_{mn}. \quad (3)$$

These polynomials are uniquely determined by $V^2(x)$. The orthonormal polynomials for an arbitrary window function $V^2(x)$ are given by

$$G_n(x) = \frac{1}{\sqrt{M_{n-1} M_n}} \begin{vmatrix} c_0 & c_1 & \cdots & c_n \\ c_1 & c_2 & \cdots & c_{n+1} \\ \vdots & \vdots & & \vdots \\ c_{n-1} & c_n & \cdots & c_{2n-1} \\ 1 & x & \cdots & x^n \end{vmatrix} \quad (4)$$

where the determinant M_n is defined by

$$M_n = |c_{i+j}|_{i,j=0,\dots,n}, \quad M_{-1} = 1 \quad (5)$$

and

$$c_n = \int_{-\infty}^{+\infty} x^n V^2(x) dx \quad (6)$$

is the n th order moment [38].

If $V^2(x)$ is even, then we can derive the following explicit expressions for the orthonormal polynomials up to order 3

$$\begin{aligned} G_0(x) &= 1/\sqrt{c_0} \\ G_1(x) &= x/\sqrt{c_2} \\ G_2(x) &= (c_0 x^2 - c_2)/\sqrt{c_0(c_0 c_4 - c_2^2)} \\ G_3(x) &= (c_2 x^3 - c_4 x)/\sqrt{c_2(c_2 c_6 - c_4^2)}. \end{aligned} \quad (7)$$

Examples of orthogonal polynomials for different window functions are listed in [39].

Under very general conditions [40] for the original signal $L(x)$, we get that

$$V(x - kT) \left[L(x) - \sum_{n=0}^{\infty} L_n(kT) \cdot G_n(x - kT) \right] = 0 \quad (8)$$

with

$$L_n(kT) = \int_{-\infty}^{+\infty} L(x) \cdot G_n(x - kT) V^2(x - kT) dx. \quad (9)$$

For instance, requiring that $L(x)$ is analytic and finite for all x is sufficient to guarantee the convergence of the series expansion in (8) for most window functions. Hence, the approximation error between a signal and a polynomial can be made arbitrarily small by taking the degree of the polynomial expansion sufficiently high, as is well known from Taylor expansions. This implies that the description of the localized signal $L(x) \cdot V(x - kT)$ can, up to an arbitrary small approximation error, be reduced to specifying a finite set of polynomial coefficients $L_n(kT)$. The signal energy within the window can be expressed in terms of the coefficients of the expansion, i.e.,

$$\int_{-\infty}^{+\infty} L^2(x) V^2(x - kT) dx = \sum_{n=0}^{\infty} L_n^2(kT). \quad (10)$$

This is the generalization of Parseval's theorem to orthonormal polynomials.

Combining (2) and (8), we get the following expansion for the complete signal

$$L(x) = \sum_{n=0}^{\infty} \sum_k L_n(kT) \cdot P_n(x - kT) \quad (11)$$

where

$$P_n(x) = G_n(x) V(x) / W(x). \quad (12)$$

Equation (9) implies that the coefficients $L_n(kT)$ can be derived from the signal $L(x)$ by convolving with the filter functions

$$D_n(x) = G_n(-x) V^2(-x) \quad (13)$$

followed by sampling at multiples of T . This mapping from the original signal $L(x)$ to the polynomial coefficients $L_n(kT)$ is called a forward polynomial transform. The signal reconstruction from these coefficients can be done according to (11) and is called an inverse polynomial transform. It consists of interpolating the coefficients $\{L_n(kT); k \text{ integer}\}$ with the pattern function $P_n(x)$ and summing over all orders n . The forward and inverse polynomial transforms are illustrated in Fig. 1.

Appendix A shows how the spectrum of the reconstructed signal is influenced if the polynomial coefficients $L_n(kT)$ are multiplied by constants t_n . This includes the case of a finite transform for which $t_n = 1$ for $0 \leq n \leq N$ and $t_n = 0$ for $n > N$. If the coefficients in the signal expansion of (11) are pointwise multiplied by fixed constants, then the output signal becomes

$$\hat{L}(x) = \sum_{n=0}^{\infty} \sum_k t_n \cdot L_n(kT) P_n(x - kT). \quad (14)$$

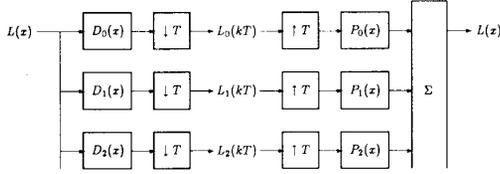


Fig. 1. Polynomial transform.

The Fourier transform of $\hat{L}(x)$ is

$$\hat{l}(\omega) = \sum_k \frac{1}{T} \sum_{n=0}^{\infty} t_n p_n(\omega) d_n \left(\omega - k \frac{2\pi}{T} \right) \cdot l \left(\omega - k \frac{2\pi}{T} \right) \quad (15)$$

where $l(\omega)$, $d_n(\omega)$, and $p_n(\omega)$ are the Fourier transforms of the respective signals $L(x)$, $D_n(x)$, and $P_n(x)$. The term for $k = 0$ represents a signal filtering with

$$h(\omega) = \frac{1}{T} \sum_{n=0}^{\infty} t_n \cdot p_n(\omega) \cdot d_n(\omega). \quad (16)$$

The aliasing terms for $k \neq 0$ arise from the sampling with period T . The filtering and aliasing effects will be discussed in the next section for the case of a finite transform and a Gaussian window. Of course, if $t_n = 1$ for all n , then the signal reconstruction is exact.

III. ONE-DIMENSIONAL HERMITE TRANSFORM

To illustrate the properties of polynomial transforms, we concentrate on the important special case that the local window function is Gaussian, i.e.,

$$V(x) = \frac{1}{\sqrt{\sqrt{\pi}\sigma}} \exp(-x^2/2\sigma^2) \quad (17)$$

where the normalization factor is such that $V^2(x)$ has unit energy. The orthogonal polynomials that are associated with $V^2(x)$ are known as the Hermite polynomials, and therefore we refer to the resulting local decomposition technique as the Hermite transform.

The reasons for focussing our attention on Gaussian windows are manifold. First, the theory is mathematically tractable in this case, so that the properties of the Hermite transform can be easily derived and evaluated. The discrete Hermite transform, to be discussed in Section VIII, is a good approximation of the analog case, and will hence have very related properties. Second, Gaussian windows which are separated by twice the spread σ are a good model for the overlapping receptive fields found in physiological experiments [15]. Third, it will turn out that the Hermite transform involves filter functions that are derivatives of Gaussians. These functions have already found widespread use in computer vision [33] and psychophysical modeling of the human visual system [34]. Hence the Hermite transform provides a broader theoretical framework for these approaches. Last but not least, the Gaussian window minimizes the product of uncertainties in the

spatial and frequency domain, which is an interesting property in image analysis [41].

In the following subsections, we will derive expressions for the different functions that play a role in the Hermite transform, i.e., the weighting function $W(x)$, the filter functions $D_n(x)$, and the pattern functions $P_n(x)$. These expressions will subsequently be used to evaluate the filtering and aliasing effects introduced by taking finite transforms.

A. Properties of the Weighting Function

Because the weighting function $W(x)$ is periodic with period T , it can be expanded into a Fourier series, i.e.,

$$W(x) = \frac{\sqrt{2\sqrt{\pi}\sigma}}{T} w(x) \quad (18)$$

with

$$w(x) = 1 + 2 \sum_{k=1}^{\infty} \exp \left[-\frac{1}{2} \left(k \frac{2\pi\sigma}{T} \right)^2 \right] \cdot \cos \left(k \frac{2\pi x}{T} \right). \quad (19)$$

This result is derived in Appendix B. The contrast of this weighting function is determined by the sampling parameter $\tau = T/\sigma$. Because we usually want to limit the number of local decompositions, especially in coding situations, it is advantageous to make τ as large as possible. On the other hand, referring to (2), we argue that $W(x)$ must be approximately constant. Otherwise, especially in a digital implementation, the postdivision by $W(x)$ would introduce a space-variant sensitivity. In Fig. 2, we have plotted the contrast of the weighting function, i.e., $(W(0) - W(T/2))/(W(0) + W(T/2))$, as a function of τ . Note that for values of τ up to 2, the window function is approximately constant.

B. Properties of the Filter Functions

The filter functions determine which information is made explicit in the coefficients of the Hermite transform. The main properties of the Hermite transform are therefore determined by these filter functions. From the general expression in (13) we derive that

$$D_n(x) = \frac{(-1)^n}{\sqrt{2^n n!}} \cdot \frac{1}{\sigma \sqrt{\pi}} H_n \left(\frac{x}{\sigma} \right) e^{-x^2/\sigma^2} \quad (20)$$

since the Hermite polynomials $\{H_n(x/\sigma); n = 0, 1, \dots\}$ are orthogonal over the Gaussian window $V^2(x)$ [39]. It is easily demonstrated [39] that the filter function $D_n(x)$ is equal to the n th order derivative of a Gaussian, i.e.,

$$D_n(x) = \frac{1}{\sqrt{2^n n!}} \cdot \frac{d^n}{d \left(\frac{x}{\sigma} \right)^n} \left[\frac{1}{\sigma \sqrt{\pi}} e^{-x^2/\sigma^2} \right]. \quad (21)$$

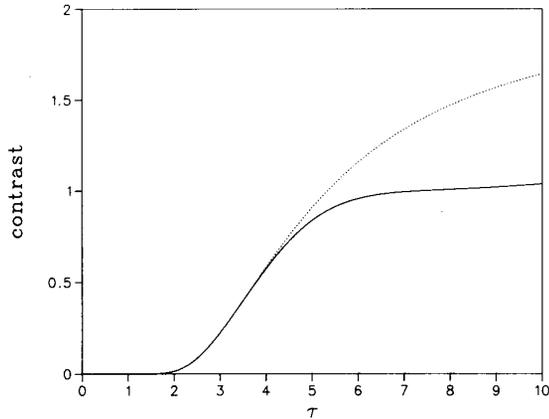


Fig. 2. Contrast of weighting function as a function of τ . The dotted line shows the contrast for the first harmonic.

The Fourier transform has a very simple expression, i.e.,

$$d_n(\omega) = \frac{1}{\sqrt{2^n n!}} \cdot (j\omega\sigma)^n e^{-(\omega\sigma)^2/4}. \quad (22)$$

It has an extreme value for $(\omega\sigma)^2 = 2n$, and hence filters of increasing order analyze successively higher frequencies in the signal. However, for large orders, the frequency peaks move very close together, so that successive filters give only very little additional information. Therefore, in practice, the Hermite transform will always be limited to a few terms. The effects of limiting the order of the Hermite transform will be discussed in the last subsection. The filter functions for $n = 0, \dots, 4$, together with their Fourier transforms, are shown in Fig. 3.

Researchers in human visual research have mainly concentrated on Gabor functions for describing the processing in the retina and cortex. The main argument for Gabor expansions is that receptive field profiles can be modeled by Gabor functions. Because receptive field profiles are only known with limited precision, however, derivatives of Gaussians can be used equally well for modeling, as has been demonstrated in some recent studies [35], [42]. They even have the advantage of requiring fewer parameters.

The past preference for using Gabor functions could be partly due to the fact that there exists a mathematical theory for the signal expansion into Gabor functions [11], [12]. This theory is, however, in contradiction with the initial assumption that receptive field profiles are Gabor functions. The fundamental problem was already recognized by Daugman [50]: "One disadvantage of the Gabor scheme is that the elementary expansion functions are not orthogonal with each other, and hence the correct code coefficients are not obtained simply by the usual inner product rule." Hence, the theory of Gabor expansions does not even lead to receptive fields! Moreover, the biorthogonal function needed for determining the coefficients of the Gabor expansion is discontinuous [12], another fact which is difficult to interpret from the point of view of visual perception.

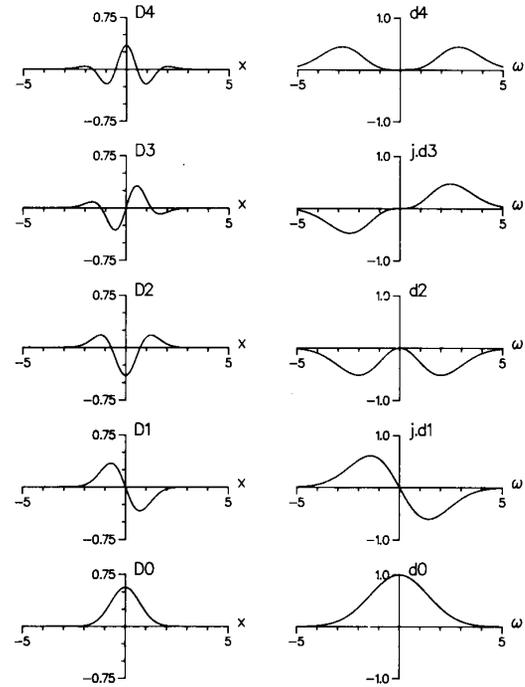


Fig. 3. Filter functions in spatial and frequency domain for $\sigma = 1$.

The Hermite transform is an alternative signal expansion technique which leads to receptive field profiles that are derivatives of Gaussian. These filter functions can be interpreted very easily. In computer vision, a large body of literature has been devoted to examining the properties of derivatives of Gaussian filters [33], [42]. With the exception of Hartley [32], the interest has been mostly in either first- or second-order derivatives, however. The theory of the Hermite transform indicates that all derivatives can be important, depending on the local signal content. Moreover, this theory also gives us an improved idea of how the outputs of different derivative operations can be used in combination. In the accompanying paper [51] we will go into this issue in more detail.

C. Properties of the Pattern Functions

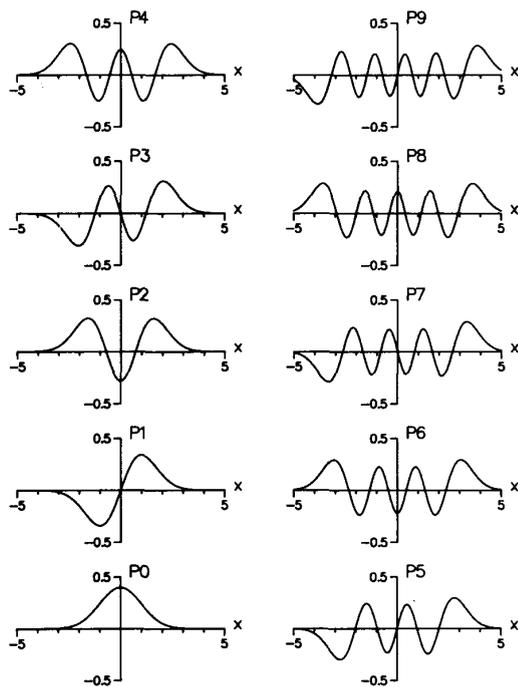
The pattern functions $P_n(x)$ are required for resynthesizing the original signal from the coefficients of the Hermite transform. They are given by the following analytical expressions

$$P_n(x) = \frac{T}{\sqrt{2^n n!}} \cdot \frac{1}{\sigma\sqrt{2\pi}} H_n\left(\frac{x}{\sigma}\right) e^{-x^2/2\sigma^2/w(x)} \quad (23)$$

where $w(x)$ is the weighting function of (19).

If $w(x) = 1$, i.e., for sampling parameter values $\tau < 2$, the pattern function $P_n(x)$ is equal to the Hermite function of order n . This implies that it is an eigenvalue of the simple harmonic oscillator problem

$$\left[\frac{x^2}{\sigma^2} - \sigma^2 \cdot \frac{d^2}{dx^2} \right] P_n(x) = (2n + 1) \cdot P_n(x). \quad (24)$$

Fig. 4. Hermite functions for $\sigma = 1$.

Moreover, the Hermite function also has the property that it is isomorphic to its Fourier transform [39], i.e.,

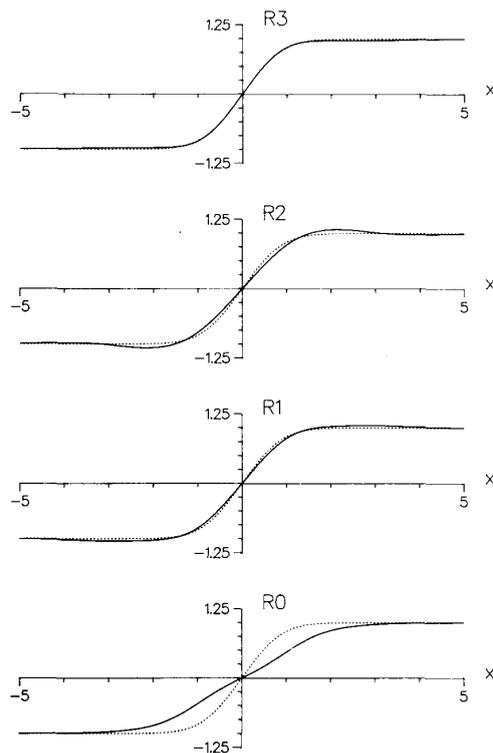
$$p_n(\omega) = \frac{T}{\sqrt{2^n n!}} \cdot (-j)^n H_n(\omega\sigma) e^{-(\omega\sigma)^2/2}. \quad (25)$$

Hence, the plots in Fig. 4 can be considered to represent either the Hermite functions themselves or their Fourier transforms. The close resemblance between these functions and truncated sines and cosines indicates that the Hermite transform is related to a harmonic analysis. The main difference with Gabor expansions is that the extreme values of a Hermite function have about equal amplitude, while Gabor functions have a Gaussian envelope.

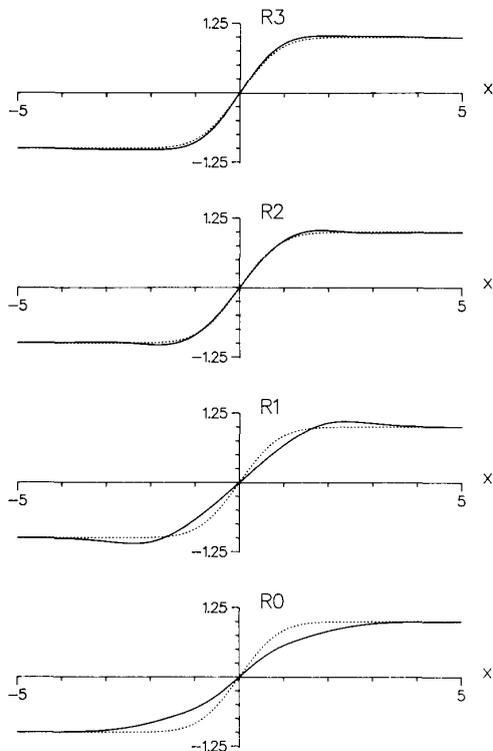
D. Finite Hermite Transform

In practice, the Hermite transform will often be limited to the first few terms. In order for the finite Hermite transform to describe the signal adequately, σ must be properly selected. This is where the problem of scale comes in, because the optimum scale σ depends on the local scene content. On the one hand, we want σ to be as large as possible because integrating over large areas improves the output signal-to-noise ratio as well as the efficiency of our signal representation. On the other hand, σ cannot be too large because then the signal cannot be described accurately by the first few terms in the Hermite expansion. The problem of selecting the right scale will be the main topic of the accompanying paper.

To get some feeling for the effects of filtering and aliasing that are introduced by the finite Hermite transform, we show some simulation results in Fig. 5. The input pat-



(a)



(b)

Fig. 5. Edge reproduction by finite Hermite transforms of order 0 to 3 for different edge positions. The dotted line is the input signal.

tern is a Gaussian edge with unity spread, i.e., $y = \text{erf}(x)$, where erf denotes the error function. The parameters of the Hermite transform are $\sigma = 1$ and $T = 2$. In Fig. 5(a), we show the output signal for increasing order of the Hermite transform in case the edge coincides with the center of an analysis window. In Fig. 5(b), we show the corresponding output signals in case the edge is located halfway between two analysis windows. These are the two extreme situations. Note that the approximation of the original signal improves with the order in both cases. The effect of aliasing in that the edge filtering is slightly dependent on the position of the edge, relative to the position of the analysis windows. This difference also decreases with increasing order of the transform.

IV. TWO-DIMENSIONAL POLYNOMIAL TRANSFORMS

The polynomial transform technique can be easily generalized to two dimensions. Given a local window function $V(x, y)$, the orthonormal polynomials $G_{m,n-m}(x, y)$, where m and $n - m$ are the degrees with respect to x and y , respectively, are uniquely determined by

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} V^2(x, y) G_{m,n-m}(x, y) G_{j,i-j}(x, y) dx dy = \delta_{ni} \delta_{mj} \quad (26)$$

for $n, i = 0, 1, \dots, \infty$; $m = 0, \dots, n$ and $j = 0, \dots, i$ [43].

The decomposition of 2D signals into localized polynomials becomes

$$L(x, y) = \sum_{n=0}^{\infty} \sum_{m=0}^n \sum_{(p,q) \in S} L_{m,n-m}(p, q) \cdot P_{m,n-m}(x - p, y - q) \quad (27)$$

where (p, q) ranges over all coordinates in a 2D sampling lattice S . The only condition for the sampling lattice is that the weighting function

$$W(x, y) = \sum_{(p,q) \in S} V(x - p, y - q) \quad (28)$$

is different from zero for all coordinates (x, y) .

The polynomial coefficients $L_{m,n-m}(p, q)$ are derived by convolving the image with the filter functions

$$D_{m,n-m}(x, y) = G_{m,n-m}(-x, -y) V^2(-x, -y) \quad (29)$$

followed by a sampling of the output at $(p, q) \in S$. The pattern functions used for interpolating the polynomial coefficients are defined by

$$P_{m,n-m}(x, y) = G_{m,n-m}(x, y) V(x, y) / W(x, y) \quad (30)$$

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

It was argued in the introduction that an image analysis should aim at decomposing a signal into patterns that are perceptually important. It has long been acknowledged, especially in computer vision and visual perception, that local 1D patterns such as edges and lines play a central role in early vision. We establish here how the best local 1D fit to an image can be found with the help of poly-

mial transforms. Using a weighted square error criterion, we minimize

$$E^2 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [K(x \cos \theta + y \sin \theta) - L(x, y)]^2 V^2(x, y) dx dy \quad (31)$$

over all 1D patterns K and angles θ .

We define the 1D window function

$$V_{\theta}^2(u) = \int_{-\infty}^{+\infty} V^2(u \cos \theta - v \sin \theta, u \sin \theta + v \cos \theta) dv \quad (32)$$

by projecting the 2D function $V^2(x, y)$ on an axis that makes an angle θ with the x axis. This window function is independent of the orientation if $V^2(x, y)$ is rotationally symmetric. We can expand the 1D pattern $K(u)$ in the basis $\{F_{n,\theta}(u); n = 0, 1, \dots\}$ of orthonormal polynomials over $V_{\theta}^2(u)$, i.e.,

$$V_{\theta}(u) \left[K(u) - \sum_{n=0}^{\infty} K_{n,\theta} \cdot F_{n,\theta}(u) \right] = 0. \quad (33)$$

Substituting the 2D and 1D polynomial expansions for $L(x, y)$ and $K(u)$, respectively, in (31), and taking the partial derivative with respect to $K_{n,\theta}$ results in the following optimum solution

$$K_{n,\theta} = \sum_{k=0}^n \sum_{l=0}^k L_{l,k-l} \cdot h_{n,\theta}(l, k - l) \quad (34)$$

for the 1D pattern coefficients, where

$$h_{n,\theta}(l, k - l) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} F_{n,\theta}(x \cos \theta + y \sin \theta) \cdot G_{l,k-l}(x, y) V^2(x, y) dx dy \quad (35)$$

is an *angle* function that is completely determined by $V^2(x, y)$.

The orthogonal polynomials $F_{n,\theta}(u)$ and the angle function can be determined without explicit knowledge of $V_{\theta}(u)$. Indeed, (4) implies that only the moments

$$c_{n,\theta} = \int_{-\infty}^{+\infty} u^n V_{\theta}^2(u) du \quad (36)$$

are needed to fully specify the orthogonal polynomials. The calculation of these moments can however be based directly on $V^2(x, y)$, since

$$c_{n,\theta} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x \cos \theta + y \sin \theta)^n V^2(x, y) dx dy. \quad (37)$$

From the orthogonality of the polynomials $F_{n,\theta}(u)$; we can derive the following properties

$$h_{n,\theta}(l, k - l) = 0 \quad \text{if } k > n \quad (38)$$

and

$$\sum_{k=0}^{\infty} \sum_{l=0}^k h_{m,\theta}(l, k-l) h_{n,\theta}(l, k-l) = \delta_{mn} \quad (39)$$

for the angle function.

The 1D approximation error

$$E^2 = \sum_{k=0}^{\infty} \sum_{l=0}^k L_{l,k-l}^2 - \sum_{n=0}^{\infty} K_{n,\theta}^2 \quad (40)$$

can be minimized over the angle θ by maximizing the directional energy

$$\sum_{n=0}^{\infty} K_{n,\theta}^2 \quad (41)$$

where $K_{n,\theta}$ is determined by the 2-D polynomial coefficients through (34). In practice, the first few terms in this directional energy measure are usually sufficient to make a good estimate for the optimum direction. Note that the directional energy is found by a simple combination of the 2D polynomial coefficients. This is computationally more efficient than using distinct filters for calculating the energy in every direction [10], [30], especially if a large number of directions are tried.

If the original image $L(x, y)$ is locally 1D, then the estimation error must be zero for the optimum angle θ , so that the 2D polynomial coefficients must satisfy

$$L_{l,k-l} = \sum_{n=k}^{\infty} K_{n,\theta} \cdot h_{n,\theta}(l, k-l). \quad (42)$$

If the image $L(x, y)$ is not locally 1D, then the latter expression can still be used as an optimal 1D approximation for the 2D polynomial coefficients.

V. TWO-DIMENSIONAL HERMITE TRANSFORM

An interesting special case of 2-D polynomial transforms arises when the window function is separable, i.e., $V(x, y) = V(x)V(y)$, and the sampling lattice is square. The filter and pattern functions are then also separable, and can hence be implemented very efficiently. For example, the polynomial coefficients are found by convolving the image with the filter functions $D_m(x)D_{n-m}(y)$, where $D_m(x)$ is the 1D filter function for window $V(x)$, followed by a sampling of the output in horizontal and vertical directions at multiples of the sample spacing T .

The Hermite transform arises if the window function is Gaussian. An important advantage of Gaussian windows in two dimensions is that they have the unique property of being both spatially separable and rotationally symmetric. The corresponding properties of the filter functions are that they are separable both spatial and polar. The Fourier transform of $D_m(x)D_{n-m}(y)$, expressed in polar coordinates $\omega_x = \omega \cos \theta$ and $\omega_y = \omega \sin \theta$, is

$$d_m(\omega_x) d_{n-m}(\omega_y) = g_{m,n-m}(\theta) \cdot d_n(\omega) \quad (43)$$

where $d_n(\omega)$ is the Fourier transform of the 1D Hermite

filter function $D_n(r)$, with r the radial coordinate, and

$$g_{m,n-m}(\theta) = \sqrt{\frac{n!}{m!(n-m)!}} \cos^m \theta \cdot \sin^{n-m} \theta \quad (44)$$

expresses the directional selectivity of the filter. Hence, filters of increasing order n analyze successively higher radial frequencies, i.e., higher spatial resolutions, similarly as in the 1D case. Filters of the same order n and different (directional) index m distinguish between different orientations in the image. The relation with the general angle function of the preceding section is

$$h_{n,\theta}(l, k-l) = \delta_{nk} \cdot g_{l,k-l}(\theta) \quad (45)$$

which results in a substantial simplification over the general case.

Daugman [45] has already demonstrated the importance of polar separable filters, i.e., filters that can be expressed as the product of a spatial frequency tuning function and an orientation tuning function. More specifically, only separable filters have the property of giving identical orientation tuning curves for different 1D patterns, such as gratings, lines, and edges. This implies that these filters can detect the orientation of a 1D pattern, independent of its internal structure.

VI. THREE-DIMENSIONAL HERMITE TRANSFORM

Although the polynomial transform technique of Section IV can be easily generalized to three dimensions, we restrict ourselves to some remarks for the case of the Hermite transform.

The three-dimensional (3D) case refers to spatiotemporal signals $L(x, y, t)$. Since we use Gaussian filters in our transform with identical spread σ along all dimensions, we have to agree on some equivalence between spatial and temporal dimensions. Therefore, we will map all spatiotemporal signals $L(x, y, t)$ onto 3D signals $L(x, y, z)$ by setting $z = u.t$. The constant u has the dimension of a velocity. In many applications, such as sampled signals, the equivalence parameter u is implicitly selected. Its choice can have far-reaching consequences, however, as it determines the velocity range to which the Hermite transform is most sensitive. There are indications that the human visual system contains two subsystems, one sensitive to high spatial and low temporal frequencies and the other sensitive to low spatial and high temporal frequencies [46]. This could be simulated by two Hermite transforms, one with small u and σ , and one with large u and σ .

The definition of the 3D Hermite transform is straightforward, i.e.,

$$L(x, y, z) = \sum_{n=0}^{\infty} \sum_{m=0}^n \sum_{l=0}^m \sum_{(p,q,r) \in S} L_{l,m-l,n-m}(p, q, r) \cdot P_l(x-p) P_{m-l}(y-q) P_{n-m}(z-r) \quad (46)$$

where the Hermite coefficients are derived from the original signal $L(x, y, z)$ by convolving with the filter func-

tions $D_l(x)D_{m-l}(y)D_{n-m}(z)$ and sampling on a lattice S . The Fourier transforms of the filter functions can be expressed in spherical coordinates $\omega_x = \omega \cos \theta \cos \phi$, $\omega_y = \omega \sin \theta \cos \phi$ and $\omega_z = \omega \sin \phi$, i.e.,

$$d_l(\omega_x)d_{m-l}(\omega_y)d_{n-m}(\omega_z) = g_{l,m-l}(\theta) \cdot g_{m,n-m}(\phi) \cdot d_n(\omega) \quad (47)$$

where $d_n(\omega)$ is the Fourier transform of the 1D Hermite filter function $D_n(r)$. The function g is identical to the one introduced in the previous section. We see that, next to directional selectivity, we now also get velocity selectivity into our filters. Indeed, the best fit of the original signal $L(x, y, z)$ by a 1D pattern

$$K((x \cos \theta + y \sin \theta) \cos \phi + z \sin \phi) \quad (48)$$

is found by maximizing the directional energy

$$\sum_{n=0}^{\infty} K_{n,\theta,\phi}^2 = \sum_{n=0}^{\infty} \left[\sum_{m=0}^n \sum_{l=0}^m g_{l,m-l}(\theta) \cdot g_{m,n-m}(\phi) L_{l,m-l,n-m} \right]^2 \quad (49)$$

over all (θ, ϕ) , provided the approximation error is weighted by $V^2(x, y, z)$, of course. If the optimum $\phi = \pi/2$ then the best 1D approximation is a purely temporal pattern. This includes, for instance, the case of uniform flicker. If the optimum $\phi \neq \pi/2$, then the best 1D fit is a pattern making an angle θ with the x axis and moving with constant velocity. The velocity vector is $(-u \tan \phi) \cdot (\cos \theta, \sin \theta)$.

The best 1D approximation of the 3D Hermite coefficients is given by

$$\hat{L}_{l,m-l,n-m} = K_{n,\theta,\phi} \cdot g_{l,m-l}(\theta) g_{m,n-m}(\phi) \quad (50)$$

for $n = 0 \dots \infty$, $m = 0 \dots n$ and $l = 0 \dots m$, with (θ, ϕ) the optimum angles.

VII. DISCRETE POLYNOMIAL TRANSFORMS

Up to now all signals and filters were assumed to be continuous. Practical applications of polynomial transforms require a formulation for discrete signals. We present two alternative ways of formulating discrete polynomial transforms.

One possible approach is to link every discrete signal to an analog one, i.e., we can restrict ourselves to analog signals

$$L(x) = \sum_q L_q \cdot I(x - q\Delta) \quad (51)$$

that are fully specified by a countable number of coefficients L_q through interpolation with $I(x)$. Applying a forward polynomial transform to this signal results in the following coefficients

$$L_n(pT) = \sum_q L_q [D_n(x) * I(x)]_{x=pT-q\Delta} \quad (52)$$

If T is a multiple of the sampling distance, i.e., $T = T_\Delta \Delta$, then the polynomial coefficients of order n are found by a discrete convolution of the sequence L_q with the filter se-

quence

$$DI_n(q) = [D_n(x) * I(x)]_{x=q\Delta} \quad (53)$$

followed by a subsampling by a factor of T_Δ . In practice, the subsampling can be combined with the filtering by only calculating part of the filtered outputs.

We can, however, also define polynomial transforms directly on discrete signals, i.e., without requiring an explicit link between analog and discrete signals. In the case of 1D polynomial transforms, the results of Section II still apply. The expressions for the weighting, filter, and pattern functions are still valid, provided we replace the continuous variable x by a discrete one. All integral expressions have to be changed into discrete sums however. For instance, (6) for the n th order moment must be replaced by

$$c_n = \sum_x x^n V^2(x) \quad (54)$$

for $n = 0, \dots, N$. If the discrete window is finite, i.e., $V(x) = 0$ for $x < N_1$ and $x > N_2$, then the polynomial transform has a finite order $N = N_2 - N_1$. Polynomial coefficients up to order N are then sufficient to get a perfect reconstruction for any discrete signal. The reason is that the discrete signal within the window $V(x)$ has only $N + 1$ degrees of freedom.

In the case of 2D polynomial transforms, most of the results for analog images can be adjusted in a straightforward way to discrete images. However, some care must be taken in applying the local 1D approximation technique of Section IV. The complication comes from the fact that in projecting the function $V^2(x, y)$ on an axis making an angle θ with the x axis, i.e.,

$$V_\theta^2(u) = \sum_v V^2(u \cos \theta - v \sin \theta, u \sin \theta + v \cos \theta) \quad (55)$$

only the sampling points have to be considered in the projection. Hence, u and v assume only the values

$$u = x \cos \theta + y \sin \theta, \quad v = -x \sin \theta + y \cos \theta \quad (56)$$

where x and y range over all (integer) values for which $V(x, y)$ is nonzero. The smoothest window functions are usually obtained by selecting the angle θ such that the projection lines pass through as many sampling points as possible, see Fig. 6.

Application of the 1D approximation technique requires knowledge of the angle function

$$h_{n,\theta}(l, k-l) = \sum_{x,y} F_{n,\theta}(x \cos \theta + y \sin \theta) \cdot G_{l,k-l}(x, y) V^2(x, y) \quad (57)$$

where $F_{n,\theta}(u)$ is the orthonormal polynomial of order n over the 1D window $V_\theta^2(u)$.

VIII. DISCRETE HERMITE TRANSFORM

In this section, we derive the discrete equivalent of the Hermite transform. It is well known that the discrete

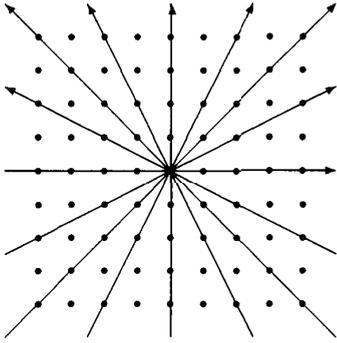


Fig. 6. Discrete angles on a 2D square sampling grid.

counterpart of a Gaussian window is a binomial window, i.e.,

$$V^2(x) = \frac{1}{2^M} C_M^x \quad (58)$$

for $x = 0, \dots, M$. The (discrete) orthonormal polynomials that are associated with this window are known as Krawtchouk's polynomials

$$G_n(x) = \frac{1}{\sqrt{C_M^n}} \sum_{k=0}^n (-1)^{n-k} C_{M-x}^{n-k} \cdot C_x^k \quad (59)$$

for $x, n = 0, \dots, M$ [38].

For large values of M , the binomial window reduces to a Gaussian window. More specifically,

$$\lim_{M \rightarrow \infty} \frac{1}{2^M} C_M^{x+(M/2)} = \frac{1}{\sqrt{\pi} \sqrt{\frac{M}{2}}} \exp \left[- \left(\frac{x}{\sqrt{\frac{M}{2}}} \right)^2 \right] \quad (60)$$

for $x = -(M/2), \dots, M/2$. It can be shown [38] that the same limiting process turns a Krawtchouk polynomial into a Hermite polynomial, i.e.,

$$\lim_{M \rightarrow \infty} G_n \left(x + \frac{M}{2} \right) = \frac{1}{\sqrt{2^n n!}} H_n \left(\frac{x}{\sqrt{\frac{M}{2}}} \right). \quad (61)$$

Hence, the discrete Hermite transform of length M approximates the analog Hermite transform of spread $\sigma = \sqrt{M/2}$. The properties of the discrete Hermite transform can therefore be predicted quite accurately from the corresponding properties of the analog Hermite transform.

We concentrate on the case that M is even. The filter and pattern functions can then be centered on the origin by shifting the binomial window over $M/2$. This leads to the following definition for the filter functions of the discrete Hermite transform

$$D_n(x) = G_n \left(\frac{M}{2} - x \right) \cdot V^2 \left(\frac{M}{2} - x \right) \quad (62)$$

for $x = -(M/2), \dots, M/2$. These functions can be expressed as

$$D_n \left(\frac{M}{2} - x \right) = \frac{(-1)^n}{2^M \sqrt{C_M^n}} \Delta^n [C_M^x \cdot C_x^n] \quad (63)$$

where

$$(-1)^n \Delta^n L(x) = \sum_{k=0}^n (-1)^k C_n^k L(x+k) \quad (64)$$

is the n th order difference operator [39]. Taking the z transform of this filter function results in

$$\begin{aligned} d_n(z) &= \sum_{x=-M/2}^{M/2} D_n(x) z^{-x} \\ &= z^{-M/2} \sqrt{C_M^n} \left(\frac{1-z}{2} \right)^n \left(\frac{1+z}{2} \right)^{M-n} \end{aligned} \quad (65)$$

or, expressed in angular frequencies

$$d_n(e^{-j\omega}) = \sqrt{C_M^n} \left(j \sin \frac{\omega}{2} \right)^n \left(\cos \frac{\omega}{2} \right)^{M-n} \quad (66)$$

for $n = 0, \dots, M$. It is obvious that for small ω this filter reduces to an n th order derivative operation, just as in the analog case.

The above filters have the important practical advantage that they can be realized by a cascade of the simple filters $z^{-1}(1+z)^2$, $z^{-1}(1-z)(1+z)$, $z^{-1}(1-z)^2$, with respective filter kernels $[1 \ 2 \ 1]$, $[-1 \ 0 \ 1]$, and $[1 \ -2 \ 1]$. Hence, with the exception of the amplification factor $\sqrt{C_M^n}$, these filters can be realized without general multiplications [47].

The calculation of the angle function $h_{n,\theta}$ is a straightforward application of (56), although the calculations may be quite lengthy for large n . Substantial deviations from the angle function of the analog Hermite transform, given in (44), will, however, only occur if M is small. In most applications, we will only be interested in the angle function $h_{n,\theta}$ for small n . Explicit expressions for $n = 0, 1, 2$ are given below

$$h_{0,\theta}(0, 0) = 1$$

$$h_{1,\theta}(0, 0) = 0$$

$$h_{1,\theta}(1, 0) = \cos \theta$$

$$h_{1,\theta}(0, 1) = \sin \theta$$

$$h_{2,\theta}(0, 0) = 0$$

$$h_{2,\theta}(1, 0) = 0$$

$$h_{2,\theta}(0, 1) = 0$$

$$h_{2,\theta}(2, 0) = \cos^2 \theta \cdot \sqrt{1 - \frac{1}{M}}$$

$$\sqrt{1 - \frac{1}{M} (\cos^4 \theta + \sin^4 \theta)}$$

$$h_{2,\theta}(1, 1) = \sqrt{2} \cos \theta \sin \theta$$

$$\sqrt{1 - \frac{1}{M} (\cos^4 \theta + \sin^4 \theta)}$$

$$h_{2,\theta}(2, 0) = \sin^2 \theta \cdot \sqrt{1 - \frac{1}{M}}$$

$$\sqrt{1 - \frac{1}{M} (\cos^4 \theta + \sin^4 \theta)}. \quad (67)$$

Note that $h_{0,\theta}$ and $h_{1,\theta}$ are the same as in the analog case. The maximum difference in $h_{2,\theta}$ between the analog and the discrete case is usually very small. For example, even for $M = 6$ it is already less than 5%.

IX. CONCLUSIONS

In the introduction we argued the importance of using smooth window functions for image analysis, the main argument being the analogy with human visual perception. One technique that applies such window functions is the Gabor expansion method. It performs local harmonic analyses on an image. In this paper we have introduced polynomial transforms, which differ from Gabor expansions because they use local polynomial decompositions. It was demonstrated that the Hermite transform, which is a polynomial transform for a Gaussian window, is in better agreement with human visual modeling than Gabor expansions.

An important advantage of polynomial transforms is that they use operators with good performance in image interpretation [48], [49]. Especially the ease of detecting local 1D patterns by means of polynomial transforms is considered important and applied repeatedly in the accompanying paper [51]. This paper will also discuss how polynomial transforms can be incorporated in a multiscale analysis. This step is necessary in order to complete the analogy with human visual perception as well as to allow a comparison with most current image coding/analysis schemes.

APPENDIX A

The filter coefficient of (9) can be written as follows:

$$L_n(kT) = \int_{-\infty}^{+\infty} l(\psi) d_n(\psi) e^{j\psi kT} \frac{d\psi}{2\pi}$$

so that the Fourier transform of (14) becomes

$$\hat{l}(\omega) = \sum_{n=0}^{\infty} t_n p_n(\omega) \int_{-\infty}^{+\infty} \frac{d\psi}{2\pi} l(\psi) d_n(\psi) \sum_k e^{-j(\omega - \psi)kT}$$

which leads directly to (15) since

$$\sum_k e^{-j(\omega - \psi)kT} = \frac{2\pi}{T} \sum_k \delta\left(\omega - \psi - k \frac{2\pi}{T}\right).$$

APPENDIX B

The Fourier transform of $w(x)$ is

$$\hat{w}(\omega) = \frac{T}{\sqrt{2\pi}\sigma} \sum_k \int_{-\infty}^{+\infty} \exp\left[-\frac{(x - kT)^2}{2\sigma^2} - j\omega x\right] dx$$

$$= T \sum_k e^{-j\omega kT} e^{-(\omega\sigma)^2/2}$$

which, using the same property as in Appendix A, leads immediately to (19).

REFERENCES

- [1] N. Nill, "A visual model weighted cosine transform for image compression and quality assessment," *IEEE Trans. Commun.*, vol. 33, pp. 551-557, June 1985.
- [2] N. Nasrabadi and R. King, "Image coding using vector quantization: A review," *IEEE Trans. Commun.*, vol. 36, pp. 957-971, Aug. 1988.
- [3] A. Jain, "Image data compression: A review," *Proc. IEEE*, vol. 69, pp. 349-389, Mar. 1981.
- [4] M. Kunt, A. Ikononopoulos, and M. Kocher, "Second generation image coding techniques," *Proc. IEEE*, vol. 73, pp. 549-574, Apr. 1985.
- [5] P. Burt and E. Adelson, "The Laplacian pyramid as a compact image code," *IEEE Trans. Commun.*, vol. 31, pp. 532-540, Apr. 1983.
- [6] M. Vetterli, "Multidimensional subband coding: Some theory and algorithm," *Signal Processing*, vol. 6, pp. 97-112, Apr. 1984.
- [7] J. Woods, "Subband coding of images," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 34, pp. 1278-1288, Oct. 1986.
- [8] P. Westerink, D. Boekee, J. Biemond, and J. Woods, "Subband coding of images using vector quantization," *IEEE Trans. Commun.*, vol. 36, pp. 713-719, June 1988.
- [9] A. Watson, "Efficiency of a model human image code," *J. Opt. Soc. Amer.*, vol. 4, pp. 2401-2417, Dec. 1987.
- [10] A. Watson, "The cortex transform: Rapid computation of simulated neural images," *Comput. Vision, Graph., Image Processing*, vol. 39, pp. 311-327, 1987.
- [11] M. Bastiaans, "Gabor's signal expansion and degrees of freedom of a signal," *Opt. Acta*, vol. 29, pp. 1223-1229, 1982.
- [12] M. Porat and Y. Zeevi, "The generalized Gabor scheme of image representation in biological and machine vision," *IEEE Trans. Pattern Analysis Mach. Intel.*, vol. 10, pp. 452-467, July 1988.
- [13] S. G. Mallat, "Multifrequency channel decompositions of images and wavelet models," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 37, no. 12, pp. 2091-2110, Dec. 1989.
- [14] S. Mallat, "A theory for multiscale signal decomposition: The scale change representation," Univ. of Pennsylvania, Dep. Comput. Inform. Sci., GRASP Lab Tech. Memo, 1987.
- [15] B. Sakitt and H. Barlow, "A model for the economical encoding of the visual image in cerebral cortex," *Biolog. Cybern.*, vol. 43, pp. 97-108, 1982.
- [16] B. Ramamurthi and A. Gersho, "Nonlinear space-variant postprocessing of block coded images," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 34, pp. 1258-1267, Oct. 1986.
- [17] W. Schreiber and D. Troxel, "Transformation between continuous and discrete representations of images: A perceptual approach," *IEEE Trans. Pattern Analysis Mach. Intel.*, vol. 7, pp. 178-186, Mar. 1985.
- [18] A. Witkin, "Scale-space filtering: A new approach to multiscale description," in *Image Understanding 1984*, 1984, ch. 3, pp. 79-95.
- [19] F. Mokhtarian and A. Mackworth, "Scale-based description and recognition of planar curves and two-dimensional shapes," *IEEE Trans. Pattern Analysis Mach. Intel.*, vol. 8, pp. 34-43, Jan. 1986.
- [20] M. Sato, T. Wada, and H. Kawarada, "A hierarchical representation of random waveforms by scale-space filtering," in *Proc. ICASSP*, 1987, pp. 273-276.
- [21] A. Yuille and T. Poggio, "Scaling theories for zero crossings," *IEEE Trans. Pattern Analysis Mach. Intel.*, vol. 8, pp. 15-25, Jan. 1986.
- [22] J. B. Martens and G. Majoor, "Scale-space coding and its perceptual relevance," *Signal Processing*, vol. 17, pp. 353-364, Aug. 1989.
- [23] E. Adelson and E. Simoncelli, "Orthogonal pyramid transforms for image coding," presented at the Fifth Workshop on Multidimensional Signal Processing, Noordwijkerhout, 1987.
- [24] J. Koenderink, "The structure of images," *Biolog. Cybern.*, vol. 50, pp. 363-370, 1984.
- [25] H. Asada and M. Brady, "The curvature primal sketch," *IEEE Trans. Pattern Analysis Mach. Intel.*, vol. 8, pp. 2-14, Jan. 1986.
- [26] P. Burt and G. van der Wal, "Iconic image analysis with the pyramid vision machine (PVM)," presented at the IEEE Workshop Comput. Architectures for Pattern Analysis, Mach. Intel., 1987.
- [27] P. Burt, C. Anderson, J. Sinniger, and G. van der Wal, "A pipelined pyramid machine," *Pyramidal Systems for Computer Vision* (NATO ASI Series). Berlin: Springer 1986, pp. 133-152.
- [28] A. Ikononopoulos, "Aspects of directional filtering and applications in image processing," in *Proc. ICASSP*, 1987, pp. 257-260.

- [29] A. Ikononopoulos and M. Kunt, "High compressing coding via directional filtering," *Signal Processing*, vol. 8, pp. 179-205, Apr. 1985.
- [30] M. Kunt, M. Bénard, and R. Leonardi, "Recent results in high-compression image coding," *IEEE Trans. Circuits Syst.*, vol. 34, pp. 1306-1336, Nov. 1987.
- [31] V. Torre and T. Poggio, "On edge detection," *IEEE Trans. Pattern Analysis Mach. Intell.*, vol. 8, pp. 147-163, Mar. 1986.
- [32] R. Hartley, "A Gaussian-weighted multiresolution edge detector," *Comput. Vision, Graph. Image Processing*, vol. 30, pp. 70-83, 1985.
- [33] J. Bevington and R. Mersereau, "Differential operator based edge and line detection," in *Proc. ICASSP*, 1987, pp. 249-252.
- [34] D. Marr and E. Hildreth, "A theory of edge detection," in *Proc. Roy. Soc. London B*, vol. 207, pp. 187-217, 1980.
- [35] R. Young, "The Gaussian derivative theory of spatial vision: Analysis of cortical cell receptive field line-weighting profiles," General Motors Res. Labs., Rep. 4920.
- [36] R. Young, "Simulation of human retinal function with the Gaussian derivative model," in *Proc. IEEE Comput. Soc. Conf. Comput. Vision Pattern Recog.*, June 1986.
- [37] R. Young, "The Gaussian derivative model for spatial vision: I. Retinal mechanisms," *Spatial Vision*, vol. 2, pp. 273-293, 1987.
- [38] G. Szegő, *Orthogonal Polynomials* (American Mathematical Society). Colloquium Publications, 1959.
- [39] M. Abramowitz and I. Stegun, *Handbook of Mathematical Functions*. New York: Dover, 1965.
- [40] R. Boas and R. Buck, *Polynomial Expansions of Analytical Functions*. New York: Springer, 1985.
- [41] R. Wilson and G. Granlund, "The uncertainty principle in image processing," *IEEE Trans. Pattern Analysis Mach. Intell.*, vol. 6, pp. 758-767, Nov. 1984.
- [42] J. Daugman, "Two-dimensional spectral analysis of cortical receptive field profiles," *Vision Res.*, vol. 20, pp. 847-856, 1980.
- [43] D. Jackson, "Formal properties of orthogonal polynomials in two variables," *Duke Math. J.*, vol. 2, pp. 423-434, 1936.
- [44] R. Haralick, "Digital step edges from zero crossing of second directional derivatives," *IEEE Trans. Pattern Analysis Machine Intell.*, vol. 6, pp. 58-68, 1984.
- [45] J. Daugman, "Six formal properties of two-dimensional anisotropic visual filters: Structural principles and frequency/orientation selectivity," *IEEE Trans. Syst., Man, Cybern.*, vol. 13, pp. 882-887, Sept./Oct. 1983.
- [46] G. Shulman and P. Mulvanny, "Discrimination of spatiotemporal patterns: The role of sustained and transient mechanisms," *Perception*, vol. 12, pp. 531-543, 1983.
- [47] W. Wells, "Efficient synthesis of Gaussian filters by cascaded uniform filters," *IEEE Trans. Pattern Analysis Mach. Intell.*, vol. 10, pp. 234-239, Mar. 1986.
- [48] C. H. Teh and R. Chin, "On image analysis by the method of moments," *IEEE Trans. Pattern Analysis Mach. Intell.*, vol. 10, pp. 496-512, July 1988.
- [49] P. Burt, "Fast algorithms for estimating local image properties," *Comput. Vision, Graph., Image Processing*, vol. 21, pp. 368-382, 1983.
- [50] J. Daugman, "Entropy reduction and decorrelation in visual coding by oriented neural receptive fields," *IEEE Trans. Biomed. Eng.*, vol. 36, pp. 107-114, Jan. 1989.
- [51] J. B. Martens, "The Hermite transform—applications," this issue, pp. 1607-1618.



Jean-Bernard Martens was born in Eeklo, Belgium, in 1956. He received the M.S. and Ph.D. degrees in electrical engineering from the University of Ghent, Belgium, in 1979 and 1983, respectively. His main research interest at that time was number theory, with applications in efficient algorithms and computer architectures for digital signal processing.

Since October 1985, he has been employed by the Institute for Perception Research in Eindhoven, The Netherlands. His current research interests lie in perceptually related image coding and processing, as well as in the mathematical modeling of the human visual system.