The Hermite Transform: A Survey

Jean-Bernard Martens

Department of Industrial Design, Eindhoven University of Technology, Den Dolech 2, P.O. Box 513, 5600 MB Eindhoven,
The Netherlands

With this survey on the Hermite transformation we want to pursue the following two goals. First, we want to provide a comprehensive and up-to-date description of the Hermite transformation, its underlying philosophy, and its most important properties and their implications for applications. As so often when publications and development go hand-in-hand, new insights have led to changes in or generalizations of already published results, and not all of these changes have been considered sufficiently substantial to be published separately. As a consequence, the existing publications on the Hermite transformation do not fully reflect our most recent insights, and the current paper intends to remedy this. Second, we also want to share some new results. Two specific new results, that is, partial signal decompositions and intersection curvatures, are therefore treated in more detail than other aspects.

1. INTRODUCTION

The Hermite transformation was introduced about 15 years agothrough two related publications [1, 2]. It was originally developed to provide a mathematical model for interpreting receptive fields in the early stages of (human) spatial vision [2, 3]. It was however also one of the first instances where an overcomplete signal representation was considered for applications. This early belief in the potential benefits of overcomplete representations, which was motivated by our belief in the superior characteristics of the human visual system, has since been shared by an increasing number of researchers. The acceptance of the Hermite transformation has grown due to demonstrated applications in diverse areas such as image coding [4–6], image fusion [7, 8], motion estimation (optic flow) [9, 10], image processing [11–15], image parameter estimation with applications in image quality prediction [16–19], image analysis [20], image indexing [21, 22], and modelling of the human visual system [23–25]. The Hermite transformation has recently also been extended to more than two dimensions using the framework of Clifford analysis [26].

The Hermite transformation is an example of an overcomplete signal representation, or signal decomposition. It is an invertible mapping of the original signal (specified by its sample values) to an alternative representation. The potential benefit of such alternative representations is that they make different characteristics of the signal explicit and hence accessible for processing or coding. The interest in a specific signal representation is therefore determined by its usefulness in applications. The recent signal-processing literature shows a fast growth, under the impulse of wavelet transformation [27, 28] and frame [29] theories, of available signal decompositions. It seems that there are endless alternatives for analyzing and resynthesizing a signal, and indeed there are. Obviously, the demand of *global correspondence*, that is, that analysis followed by synthesis results in the original signal, is not very restrictive. There is therefore ample room for imposing additional demands on signal decompositions. In this paper, we specifically consider signal decompositions from the point of view of applications in adaptive image processing (and coding). We argue that these applications are greatly simplified if signal decompositions have the additional property of *local correspondence*, that is, if signals can be locally recovered from the decomposition, using only the coefficients at the position where processing is required. The Hermite transformation is a specific example of such a local signal decomposition, but one with some very useful extra features.

The adaptive processing (or coding) of images with the help of local image decompositions will be a three-stage process, as shown in Figure 1 [20]. In the analysis stage *A* the image is decomposed into a sum of a priori selected (basic) functions. We assume that these functions span the signal space. They may comprise either a basis, in case they are linearly independent, or a frame, in case they are linearly dependent. The analysis stage is equivalent to a transformation from an image to a set of coefficients, that is, the weights that need to be assigned to the basic functions in order to reconstruct the image. We assume here that all these basic functions are displaced versions of a limited set of templates of
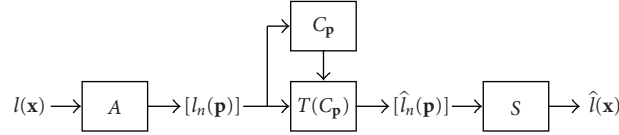
FIGURE 1: Algorithmic structure for adaptive image processing with local image decompositions. In the analysis step A, the input image $l(\mathbf{x})$ is transformed into a set of coefficients $[l_n(\mathbf{p})]$, where n indexes the different templates and $\mathbf{p}$ the different positions. The coefficients $[l]_\mathbf{p} = [l_n(\mathbf{p}),\ \forall n]$ at a given position $\mathbf{p}$ undergo a transformation $T(C_\mathbf{p})$. The control variables $C_\mathbf{p}$ at position $\mathbf{p}$ are derived from $[l]_\mathbf{p}$ and are used to select the desired transformation T. The transformed coefficients $[\hat{l}_n(\mathbf{p})]$ are combined in the synthesis step S to obtain the processed image $\hat{l}(\mathbf{x})$.

approximately the same size. The extension to basic functions of substantially different sizes can be accomplished by extending the signal decomposition to a multiresolution scheme, that is, by incorporating it in a pyramid structure [1, 11, 28, 30] or in a scale-space setting [31–34]. The weights of the basic functions can be assembled according to template (i.e., the weights for all positions of a given template) or according to position (i.e., the weights for all templates at a given position). In the second stage of the adaptive algorithm of Figure 1, the coefficients are assembled according to position. The original coefficients at a given position are interpreted and mapped (by T) into processed coefficients. Since only coefficients at the same position are combined, this mapping can easily be made to vary across positions. To implement this adaptivity, we derive at each position one or more control variables $C_\mathbf{p}$ that determine which (possibly nonlinear) transformation T has to be applied to the coefficients at that position. These control variables are intended to measure changes in the image itself (e.g., uniform versus edge region, or variable orientation), or changes in the degradation (e.g., level-dependent noise such as Poisson noise), so that an adequate transformation can be selected. An image decomposition that allows the image to be reconstructed locally obviously allows any desired local property to be derived. Some important local features that will be introduced later are the local average, the local energy (or contrast), the local dimensionality, and the local orientation. In the synthesis stage S of the adaptive processing algorithm, the processed coefficients are applied to their respective basic functions. The weighted sum is the processed output image.

In Section 2, we formalize the concept of local signal decompositions. We show how the idea of performing identical measurements at regular intervals leads to the definition of filters and filter banks. Next, in Section 3, we introduce partial signal decompositions as a specific case of local signal decompositions where the number of required filter operations can be substantially reduced. It is shown in Section 4 how the simplest instance of such a partial signal decomposition, called residue-image processing [15], can be applied for noise reduction and contrast enhancement. Starting from Section 5, we focus on one specific case of local decompositions, that is, the Hermite transformation for two-dimensional signals [1, 2, 13]. We define the Hermite transformation and show how the mapping between the image and the Hermite coefficients can be implemented. We also

discuss some important properties of the Hermite transformation. Next, in Section 6, we show how local coordinate axes that are oriented along an image-dependent direction can be created. Many geometric image features, such as the isophote and flow-line curvatures and the newly-introduced intersection curvatures, are most easily expressed in a coordinate system, that is, oriented along the gradient direction. These geometric features are discussed in Section 7, and it is shown how they can be used to identify conspicuous details (such as corners, junctions, extrema, etc.) in images. Such details are for instance well suited for performing image matching, such as that required for stereo vision and motion estimation. In Section 8 we apply the concept of partial signal decompositions to the Hermite transformation in order to devise an anisotropic noise-reduction algorithm.

2. LOCAL SIGNAL DECOMPOSITIONS

We assume that the input signals $l(\mathbf{x})$ are defined for a (compact) subset $\mathbb{F}$ of the D-dimensional Euclidean space $\mathbb{R}^D$. This subset $\mathbb{F}$ can be either continuous or discrete. In the case of image processing, the dimension $D = 2$, for static images, or $D = 3$, for image sequences. We assume that the image analysis is linear and hence consists of bounded linear functionals [24], that is, linear mappings from the input signal $l(\mathbf{x})$ to the finite real/complex numbers that can be expressed as the inproduct between a test function $\phi(\mathbf{x})$ and the input function $l(\mathbf{x})$. We moreover assume that this inproduct can be expressed as an integral

$$\langle \phi(\mathbf{x}), l(\mathbf{x}) \rangle = \int_\mathbb{F} \phi(\mathbf{x}) \cdot l(\mathbf{x}) d\mathbf{x}, \tag{1}$$

in case of continuous signals, or as a sum

$$\langle \phi(\mathbf{x}), l(\mathbf{x}) \rangle = \sum_\mathbb{F} \phi(\mathbf{x}) \cdot l(\mathbf{x}), \tag{2}$$

in case of discrete signals. The space of test functions that we consider for analyzing the image is denoted by V. Test functions are very often Schwartz functions (smooth localized functions) [35]. The signal space contains those signals that result in finite measurements for all test functions in V, and is hence equal to the dual space V' of bounded linear functionals on V.

In a local signal decomposition, we decompose a D-dimensional input signal $l(\mathbf{x})$, for $\mathbf{x} \in \mathbb{F}$, into windowed signals $l(\mathbf{x}) \cdot w(\mathbf{x} - \mathbf{p})$, for $\mathbf{p} \in \mathcal{P}$. The windowing function
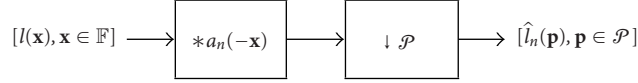
FIGURE 2: Mapping from the input image $[l(\mathbf{x}),\ \mathbf{x} \in \mathbb{F}]$ to the coefficients $[\hat{l}_n(\mathbf{p}),\ \mathbf{p} \in \mathcal{P}]$ of order n, depicted as a two-stage process. The standard notations $*$ and $\downarrow$ are used to denote the convolution and (down-)sampling operations, respectively.

$w(\mathbf{x})$ has limited support, so that $w(\mathbf{x} - \mathbf{p})$ is zero except in the neighbourhood of sampling position $\mathbf{p} \in \mathcal{P}$. The windowed signal $l(\mathbf{x}) \cdot w(\mathbf{x} - \mathbf{p})$ is therefore localized around position $\mathbf{p} \in \mathcal{P}$. The set $\mathcal{P}$ contains all positions $\mathbf{p}$ on a regular lattice Λ [36] for which $w(\mathbf{x} - \mathbf{p})$ overlaps with $\mathbb{F}$, that is, $w(\mathbf{x} - \mathbf{p}) \neq 0$ for at least one coordinate $\mathbf{x} \in \mathbb{F}$.

Each of the windowed signals $l(\mathbf{x}) \cdot w(\mathbf{x} - \mathbf{p})$ can be decomposed into a sum of orthonormal basis functions $\varphi_n(\mathbf{x} - \mathbf{p}) \cdot w(\mathbf{x} - \mathbf{p})$, $n \in \mathcal{N}$, such that

$$w(\mathbf{x} - \mathbf{p}) \left[l(\mathbf{x}) - \sum_{n \in \mathcal{N}} l_n(\mathbf{p}) \cdot \varphi_n(\mathbf{x} - \mathbf{p}) \right] = 0, \quad (3)$$

where the coefficients are obtained via

$$l_n(\mathbf{p}) = \langle a_n(\mathbf{x} - \mathbf{p}), l(\mathbf{x}) \rangle = \int_{\mathbb{F}} a_n(\mathbf{x} - \mathbf{p}) \cdot l(\mathbf{x}) d\mathbf{x}, \quad (4)$$

for $n \in \mathcal{N}$ and $\mathbf{p} \in \mathcal{P}$ (equality in (3) actually means convergence in the mean-square sense) [20, 24]. The (discrete) set $\mathcal{N}$ indexes the basis functions. The function

$$a_n(\mathbf{x}) = \varphi_n(\mathbf{x}) \cdot w^2(\mathbf{x}) \quad (5)$$

is referred to as the *analysis function* of order n. The mapping between the original signal $l(\mathbf{x})$ and the coefficients $[l_n(\mathbf{p})]$ for all orders n and positions $\mathbf{p}$ specifies the analysis stage A of Figure 1.

If measurements are performed at all positions, then (4), taken over all positions $\mathbf{p} \in \mathbb{R}^D$, defines a D-dimensional *convolution* or *filtering* between the input signal $l(\mathbf{x})$ and the filter with impulse response $a_n(-\mathbf{x})$, or equivalently, a *correlation* between the input signal $l(\mathbf{x})$ and the analysis function $a_n(\mathbf{x})$. Only a limited number of these output values are however required, since $\mathcal{P} \subset \mathbb{R}^D$. This operation of selecting a subset of values is called (down-)*sampling*. The combined operation is depicted in Figure 2. The complete signal analysis A contains several of these filter/sampling combinations, that is, one for each index $n \in \mathcal{N}$, and is called a *multirate filter bank*. The term multirate refers to the fact that the domain $\mathbb{F}$ on which the original signal is defined is usually different from the domain $\mathcal{P}$ on which the filter coefficients are required.

There is no unique way of reconstructing the signal $l(\mathbf{x})$ from the coefficients $[l_n(\mathbf{p}); n \in \mathcal{N}, \mathbf{p} \in \mathcal{P}]$ since any weighted summation

$$l(\mathbf{x}) = \sum_{\mathbf{p} \in \mathcal{P}} \sum_{n \in \mathcal{N}} l_n(\mathbf{p}) \cdot r_n(\mathbf{x} - \mathbf{p}), \quad (6)$$

with

$$r_n(\mathbf{x} - \mathbf{p}) = \varphi_n(\mathbf{x} - \mathbf{p}) \frac{w(\mathbf{x} - \mathbf{p}) u(\mathbf{x} - \mathbf{p})}{\sum_{\mathbf{q} \in \mathcal{P}} w(\mathbf{x} - \mathbf{q}) u(\mathbf{x} - \mathbf{q})}, \quad (7)$$

for which

$$\sum_{\mathbf{p} \in \mathcal{P}} w(\mathbf{x} - \mathbf{p}) u(\mathbf{x} - \mathbf{p}) \neq 0, \quad (8)$$

for all argument values $\mathbf{x} \in \mathbb{F}$, is for instance valid. The reason why there is no unique reconstruction is that the coefficients of the local signal decomposition are not independent if the window functions $w(\mathbf{x} - \mathbf{p})$ at neighboring positions $\mathbf{p}$ are overlapping. This means that, although there is a unique set of coefficients $[l_n(\mathbf{p}); n \in \mathcal{N}, \mathbf{p} \in \mathcal{P}]$ for any signal $l(\mathbf{x})$, there are many sets of coefficients that do not correspond to a signal. Hence, the set of physically realizable coefficients is a subspace of the space of all coefficients.

To arrive at a unique solution for $u(\mathbf{x})$, we have to introduce an additional condition for the reconstruction. One obvious condition is described in the mathematical theory of frames [29]. If we reconstruct a signal from a set of coefficients that are not physically realizable, and subsequently analyze this signal, then we require the output coefficients to be the closest realizable set of coefficients. In mathematical terms, we impose that synthesis S followed by analysis A is equivalent to an orthogonal projection $P = A \cdot S$ onto the set of realizable coefficients. This is a useful property in our application since the processed coefficients in adaptive signal processing cannot be guaranteed to be realizable; it is therefore important to design the synthesis stage such that the coefficients of the output signal are automatically mapped onto the closest realizable coefficients.

To find the unique reconstruction (synthesis) satisfying the above condition for the given set of analysis functions, we have to determine the frame operator F_a which maps any signal $l(\mathbf{x})$ to the corresponding signal

$$F_a l(\mathbf{x}) = \sum_{\mathbf{p} \in \mathcal{P}} \sum_{n \in \mathcal{N}} l_n(\mathbf{p}) \cdot a_n(\mathbf{x} - \mathbf{p}), \quad (9)$$

for $\mathbf{x} \in \mathbb{F}$, that arises by also using the analysis functions to perform the signal reconstruction [29]. In the case under consideration, the frame operator reduces to a multiplication by a fixed function, namely,

$$F_a l(\mathbf{x}) = l(\mathbf{x}) \cdot \sum_{\mathbf{p} \in \mathcal{P}} w^2(\mathbf{x} - \mathbf{p}), \quad (10)$$

for $\mathbf{x} \in \mathbb{F}$. The inverse frame operator exists if

$$h^2(\mathbf{x}) = \sum_{\mathbf{p} \in \mathcal{P}} w^2(\mathbf{x} - \mathbf{p}) \neq 0, \quad (11)$$

for all $\mathbf{x} \in \mathbb{F}$. Mapping the analysis functions through this inverse frame operator provides the required *synthesis functions*

$$s_n(\mathbf{x}) = \varphi_n(\mathbf{x}) \cdot v^2(\mathbf{x}), \quad (12)$$
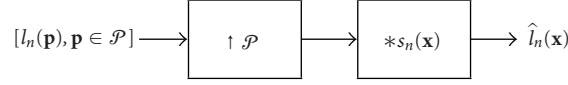
FIGURE 3: Mapping from the coefficients $[l_n(\mathbf{p}), \mathbf{p} \in \mathcal{P}]$ to the interpolated image $\hat{l}_n(\mathbf{x})$ of order n, depicted as a two-stage process. The standard notations $*$ and $\uparrow$ are used to denote the convolution and upsampling operations, respectively.
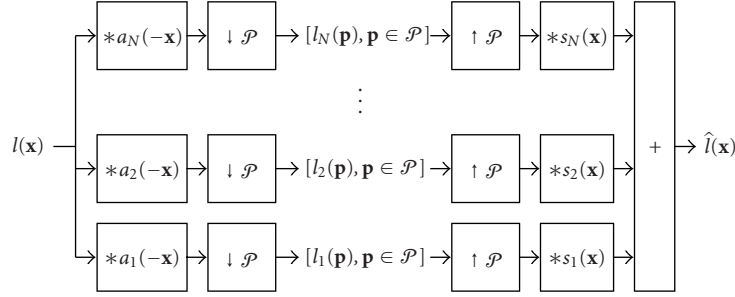

FIGURE 4: Local signal analysis and synthesis, where N denotes the number of different templates (N may be infinite).

for $n \in \mathcal{N}$, where

$$v^2(\mathbf{x}) = \frac{w^2(\mathbf{x})}{h^2(\mathbf{x})} = \frac{w^2(\mathbf{x})}{\sum_{\mathbf{p} \in \mathcal{P}} w^2(\mathbf{x} - \mathbf{p})} \tag{13}$$

is a modified window. This latter window has the property that the sum of displaced copies $v^2(\mathbf{x} - \mathbf{p})$ for all positions $\mathbf{p}$ on the sampling lattice $\mathcal{P}$ is equal to one for all positions $\mathbf{x} \in \mathbb{F}$.

The resulting synthesis stage S is specified by

$$\begin{aligned} l(\mathbf{x}) &= \sum_{\mathbf{p} \in \mathcal{P}} \sum_{n \in \mathcal{N}} l_n(\mathbf{p}) \cdot s_n(\mathbf{x} - \mathbf{p}) \\ &= \frac{\sum_{\mathbf{p} \in \mathcal{P}} \left[\sum_{n \in \mathcal{N}} l_n(\mathbf{p}) \cdot a_n(\mathbf{x} - \mathbf{p}) \right]}{\sum_{\mathbf{p} \in \mathcal{P}} w^2(\mathbf{x} - \mathbf{p})}, \end{aligned} \tag{14}$$

for $\mathbf{x} \in \mathbb{F}$. This optimal synthesis corresponds to $u(\mathbf{x}) = w(\mathbf{x})$ in our earlier notation.

If an impulse signal with unit weight at position $\mathbf{p}$ is input into a filter with impulse response $s_n(\mathbf{x})$, then the output is, by definition, equal to $s_n(\mathbf{x} - \mathbf{p})$. This implies that the signal

$$l_n(\mathbf{x}) = \sum_{\mathbf{p} \in \mathcal{P}} l_n(\mathbf{p}) \cdot s_n(\mathbf{x} - \mathbf{p}) \tag{15}$$

is the output of this filter in case the input is an array of impulses, where the weight of the impulse at position $\mathbf{p} \in \mathcal{P}$ is equal to the coefficient $l_n(\mathbf{p})$. This operation of creating an array of impulses from a set of coefficients is called *upsampling*. The combined operation of upsampling and filtering is called *interpolation* and is depicted in Figure 3. The complete signal synthesis S contains several of these interpolations, that is, one for each index $n \in \mathcal{N}$, and is hence a multirate filter bank. The synthesized signal arises by combining all outputs of this multirate filter bank.

The overall image analysis and synthesis is depicted in Figure 4. In case of Figure 4, the coefficients $[l_n(\mathbf{p})]$ that are derived from the original image are not modified, so that the output signal satisfies $\hat{l}(\mathbf{x}) = l(\mathbf{x})$, for $\mathbf{x} \in \mathbb{F}$. If $\mathbb{F}$ is discrete, then the local signal decomposition can be used to interpolate the input signal, that is, to construct signal values $\hat{l}(\mathbf{x})$, for $\mathbf{x} \notin \mathbb{F}$ [13]. In case of image processing or coding, the original coefficients are mapped to modified coefficients $[\hat{l}_n(\mathbf{p})]$ before entering the synthesis stage [1], so that $\hat{l}(\mathbf{x}) \neq l(\mathbf{x})$, for $\mathbf{x} \in \mathbb{F}$.

Note that a local signal decomposition satisfies not only the global correspondence of (14), that is, synthesis after analysis results in the original signal, but also the local correspondences of (3). The latter correspondences imply that the local description $[l]_{\mathbf{p}} = [l_n(\mathbf{p}), n \in \mathcal{N}]$ completely specifies the signal $l(\mathbf{x})$ within the local window $w(\mathbf{x} - \mathbf{p})$ at position $\mathbf{p}$. The windowed signal

$$l_w(\mathbf{x} - \mathbf{p}) = w^2(\mathbf{x} - \mathbf{p}) \cdot l(\mathbf{x}) = \sum_{n \in \mathcal{N}} l_n(\mathbf{p}) \cdot a_n(\mathbf{x} - \mathbf{p}) \tag{16}$$

also expresses the contribution of the coefficients $[l]_{\mathbf{p}} = [l_n(\mathbf{p}), n \in \mathcal{N}]$ at position $\mathbf{p} \in \mathcal{P}$ to the overall synthesized signal, in a way that is specified in (14).

3. PARTIAL SIGNAL DECOMPOSITIONS

One obvious disadvantage of local signal decompositions is that a large number of filter operations (i.e., equal to the cardinality of the index set $\mathcal{N}$) may be required in the most general case in order to create the processed signal

$$\begin{aligned} \hat{l}(\mathbf{x}) &= \sum_{\mathbf{p} \in \mathcal{P}} \sum_{n \in \mathcal{N}} \hat{l}_n(\mathbf{p}) \cdot s_n(\mathbf{x} - \mathbf{p}) \\ &= \frac{\sum_{\mathbf{p} \in \mathcal{P}} \left[\sum_{n \in \mathcal{N}} \hat{l}_n(\mathbf{p}) \cdot a_n(\mathbf{x} - \mathbf{p}) \right]}{\sum_{\mathbf{p} \in \mathcal{P}} w^2(\mathbf{x} - \mathbf{p})}. \end{aligned} \tag{17}$$

The specific dependency of the output coefficients $[\hat{l}]_{\mathbf{p}} = [\hat{l}_n(\mathbf{p}), n \in \mathcal{N}]$ at position $\mathbf{p}$ on the corresponding coefficients $[l]_{\mathbf{p}} = [l_n(\mathbf{p}), n \in \mathcal{N}]$ of the input signal may however

allow for substantial simplifications. We discuss an important special case.

We assume that the index set can be subdivided in two subsets, that is, $\mathcal{N} = \mathcal{N}_o \cup \mathcal{N}_r$, in such a way that the output coefficients $[\widehat{l}_n(\mathbf{p}), \ n \in \mathcal{N}_o]$ at position $\mathbf{p}$ only depend on the corresponding input coefficients $[l_n(\mathbf{p}), \ n \in \mathcal{N}_o]$, and the remaining coefficients at position $\mathbf{p}$ are all transformed in an identical way, that is,

$$\widehat{l}_n(\mathbf{p}) = \kappa(\mathbf{p}) \cdot l_n(\mathbf{p}), \qquad (18)$$

for $n \in \mathcal{N}_r$. We moreover assume that the adaptive weighting factor $\kappa(\mathbf{p})$ on the coefficients in the index set $\mathcal{N}_r$ can be derived from $[l_n(\mathbf{p}), \ n \in \mathcal{N}_o]$ and the local energy

$$E(\mathbf{p}) = \int_{\mathbb{F}} w^2(\mathbf{x} - \mathbf{p}) \cdot l^2(\mathbf{x})d\mathbf{x}. \qquad (19)$$

Because the local basis $\{\varphi_n(\mathbf{x} - \mathbf{p}), \ n \in \mathcal{N}\}$ was assumed to be orthonormal, this local energy can be split into

$$E_o(\mathbf{p}) = \sum_{n \in \mathcal{N}_o} \left| l_n(\mathbf{p}) \right|^2, \qquad (20)$$

the local energy in the signal component

$$\sum_{n \in \mathcal{N}_o} l_n(\mathbf{p}) \cdot \varphi_n(\mathbf{x} - \mathbf{p})w(\mathbf{x} - \mathbf{p}) \qquad (21)$$

described by the first index set $\mathcal{N}_o$, and the *residue energy* $E_r(\mathbf{p}) = E(\mathbf{p}) - E_o(\mathbf{p})$, that is, the local energy in the *residue signal*

$$\sum_{n \in \mathcal{N}_r} l_n(\mathbf{p}) \cdot \varphi_n(\mathbf{x} - \mathbf{p})w(\mathbf{x} - \mathbf{p})$$
$$= \left[l(x) - \sum_{n \in \mathcal{N}_o} l_n(\mathbf{p}) \cdot \varphi_n(\mathbf{x} - \mathbf{p}) \right] w(\mathbf{x} - \mathbf{p}) \qquad (22)$$

that is described by the second index set $\mathcal{N}_r$.

Using the above assumptions, it is easily verified that the processed signal can be expressed as

$$\widehat{l}(\mathbf{x}) = \sum_{\mathbf{p} \in \mathcal{P}} \sum_{n \in \mathcal{N}_o} [\widehat{l}_n(\mathbf{p}) - \kappa(\mathbf{p}) \cdot l_n(\mathbf{p})] \cdot s_n(\mathbf{x} - \mathbf{p}) + l(\mathbf{x}) \cdot \kappa(\mathbf{x}). \qquad (23)$$

This so-called *partial signal decomposition* leads to the modified algorithm depicted in Figure 5. Only a partial transformation involving the indices $n \in \mathcal{N}_o$ is hence required in this case to determine the processed signal. The missing information for the other indices $n \in \mathcal{N}_r$ can be recovered from the original signal. The weighting function

$$\kappa(\mathbf{x}) = \sum_{\mathbf{p} \in \mathcal{P}} \kappa(\mathbf{p}) \cdot v^2(\mathbf{x} - \mathbf{p}) \qquad (24)$$

that must be applied to the original signal in order to accomplish this correction is obtained by interpolating the adaptive factors $[\kappa(\mathbf{p}), \ \mathbf{p} \in \mathcal{P}]$ with a filter with impulse response equal to the modified window function $v^2(\mathbf{x})$. In case of image coding, where the original signal is not available at the receiver side, we obviously need to choose $\kappa(\mathbf{p}) = 0, \ \forall \mathbf{p} \in \mathcal{P}$. Nonzero values of $\kappa(\mathbf{p})$ are however useful in case of image processing.

4. RESIDUE-IMAGE PROCESSING

4.1. Algorithm

The simplest case of a partial signal decomposition arises when the index set $\mathcal{N}_o$ contains only a single component, that is, the local average signal value

$$l_o(\mathbf{p}) = \int_{\mathbb{F}} w^2(\mathbf{x} - \mathbf{p}) \cdot l(\mathbf{x})d\mathbf{x}, \qquad (25)$$

and the residue signal $l(\mathbf{x}) - l_o(\mathbf{p})$ is the local deviation from this average value. If we assume that the window function is normalized, that is, if $\int_{\mathbb{F}} w^2(\mathbf{x})d\mathbf{x} = 1$, then this corresponds to a single orthonormal basis vector $\varphi_o(\mathbf{x}) = 1$ in the index set $\mathcal{N}_o$. If we impose that the local average signal value $l_o(\mathbf{p})$ is not altered by the processing, then the partial signal decomposition can be expressed as

$$\widehat{l}(\mathbf{x}) = \sum_{\mathbf{p} \in \mathcal{P}} v^2(\mathbf{x} - \mathbf{p}) \cdot \{l_o(\mathbf{p}) + \kappa(\mathbf{p}) \cdot [l(\mathbf{x}) - l_o(\mathbf{p})]\}$$
$$= \sum_{\mathbf{p} \in \mathcal{P}} v^2(\mathbf{x} - \mathbf{p}) \cdot [1 - \kappa(\mathbf{p})]l_o(\mathbf{p}) + l(\mathbf{x}) \cdot \kappa(\mathbf{x}), \qquad (26)$$

where $\kappa(\mathbf{p})$ denotes the residue amplification factor at position $\mathbf{p} \in \mathcal{P}$, and $\kappa(\mathbf{x})$ is the corresponding interpolated function.

In uniform regions of the image, the residue energy

$$E_r(\mathbf{p}) = \int_{\mathbb{F}} w^2(\mathbf{x} - \mathbf{p}) \cdot [l(\mathbf{x}) - l_o(\mathbf{p})]^2 d\mathbf{x}$$
$$= \int_{\mathbb{F}} w^2(\mathbf{x} - \mathbf{p}) \cdot l^2(\mathbf{x})d\mathbf{x} - l_o^2(\mathbf{p}) \qquad (27)$$

will be small, while in transition regions (i.e., near edges) $E_r(\mathbf{p})$ will be large. This residue energy can hence be used as a control variable in an adaptive processing algorithm that requires different processing for uniform and transition regions. This residue energy can be derived quite efficiently, since it involves only two filter operations, one on the original signal $l(\mathbf{x})$, and one on the squared signal $l^2(\mathbf{x})$.

Both noise reduction and contrast enhancement can be achieved using residue-image processing [15]. The relationship between the residue amplification factor $\kappa(\mathbf{p})$ and the residue energy $E_r(\mathbf{p})$, or *residue amplitude* $A_r(\mathbf{p}) = \sqrt{E_r(\mathbf{p})}$, makes the adopted strategy explicit. Many different approaches can be imagined, but we discuss only one possibility.

If the residue amplitude is large, say $A_r(\mathbf{p}) > A_v$, then we can assume that the contrast is sufficiently high to guarantee clear visibility, and that no enhancement is required, that is, $\kappa(\mathbf{p}) = 1$ for $A_r(\mathbf{p}) > A_v$. If, on the contrary, the residue amplitude is small, say $A_r(\mathbf{p}) < A_t$, then the residue signal may be assumed to be noise on a uniform background, and this noise can be removed by setting $\kappa(\mathbf{p}) = 0$ for $A_r(\mathbf{p}) < A_t$. For $A_t \leq A_r(\mathbf{p}) \leq A_v$, the residue signal is judged to contain significant information, but needs to be amplified to become clearly visible. An obvious condition for the processing is that the output amplitude $\kappa(\mathbf{p}) \cdot A_r(\mathbf{p})$ should be an increasing
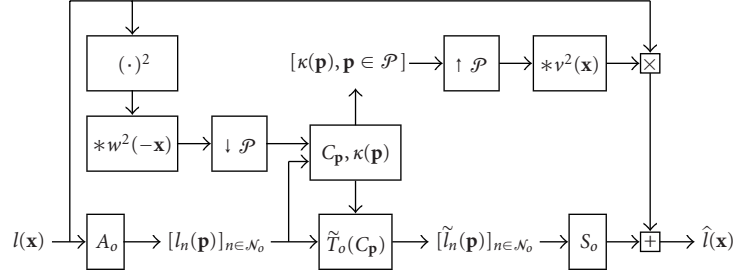
FIGURE 5: Algorithmic structure for partial image decompositions. In the analysis step A_o, the input image $l(\mathbf{x})$ is transformed into a limited set of coefficients. These coefficients $[l_n(\mathbf{p}),\ n \in \mathcal{N}_o]$ at a given position $\mathbf{p}$ undergo the transformation $\widetilde{T}_o(C_\mathbf{p}) = T_o(C_\mathbf{p}) - \kappa(\mathbf{p}) \cdot I_o$, where $T_o(C_\mathbf{p})$ is the intended transformation and I_o denotes the identity transformation. The control variables $C_\mathbf{p}$ at position $\mathbf{p}$, which are used to select the desired transformation $T_o(C_\mathbf{p})$ and the weighting factor $\kappa(\mathbf{p})$ on the missing coefficients, can be derived from $[l_n(\mathbf{p}),\ n \in \mathcal{N}_o]$ and the local energy $E(\mathbf{p})$. The transformed coefficients $[\widetilde{l}_n(\mathbf{p}) = \widehat{l}_n(\mathbf{p}) - \kappa(\mathbf{p}) \cdot l_n(\mathbf{p}),\ n \in \mathcal{N}_o]$, where $[\widehat{l}_n(\mathbf{p}),\ n \in \mathcal{N}_o]$ are the intended coefficients for the processed image $\widehat{l}(\mathbf{x})$, are combined in the synthesis step S_o to obtain part of the output image. The missing part of the processed image is extracted from the original image $l(\mathbf{x})$ by a weighting function $\kappa(\mathbf{x})$ that interpolates the factors $[\kappa(\mathbf{p}),\ \mathbf{p} \in \mathcal{P}]$ to all positions $\mathbf{x} \in \mathbb{F}$.

function of the input amplitude $A_r(\mathbf{p})$. Many functions satisfy this condition, and a simple example is

$$\kappa(A_r(\mathbf{p})) \cdot A_r(\mathbf{p}) = s \cdot A_r(\mathbf{p}) + (1 - s) \cdot A_v \qquad (28)$$

for $A_t \le A_r(\mathbf{p}) \le A_v$, where $0 \le s \le 1$ denotes the slope of the input-output amplitude characteristic. In this case, the maximum value of the amplification factor is

$$\kappa(A_t) = s + (1 - s) \cdot \frac{A_v}{A_t} \qquad (29)$$

for $A_r(\mathbf{p}) = A_t$, and decreases to 1 for $A_r(\mathbf{p}) = A_v$. Alternative expressions for $\kappa(A_r(\mathbf{p}))$ that avoid the discontinuity at $A_r(\mathbf{p}) = A_t$ may also be considered.

In Figure 6, we illustrate residue-image processing on an 8-bit subtractive angiography image (values between 0 and 255). The threshold value $A_t = 3$ was derived from the residue amplitude histogram (i.e., $A_t = 2 \cdot A_m$, where A_m is the mode of the residue amplitude histogram, see next section), and the remaining parameters were $A_v = 15$ and $s = 0.1$. A separable binomial window function [2] of size 7 (i.e., with coefficients $w^2 = [1\ 6\ 15\ 20\ 15\ 6\ 1]/64$ in one dimension) was used as the window function.

4.2. Energy histograms and noise estimation

Since images usually contain noise, the residue energy $E_r(\mathbf{p})$ will not most often be exactly zero, not even for regions that we consider to be uniform. We hence have to decide on a threshold value $E_t = A_t^2$ for the residue energy measure $E_r(\mathbf{p})$, below which image positions are classified as belonging to uniform regions. As part of this process, we need to estimate noise characteristics. This noise estimation may also be interesting in its own right, for example, to objectively characterize the quality of the image acquisition [17, 19, 24].

It has been demonstrated [15, 24] that the first-order statistic, that is, the histogram or probability density function (PDF), of $E_r(\mathbf{p})$, at least in uniform regions of the image, can
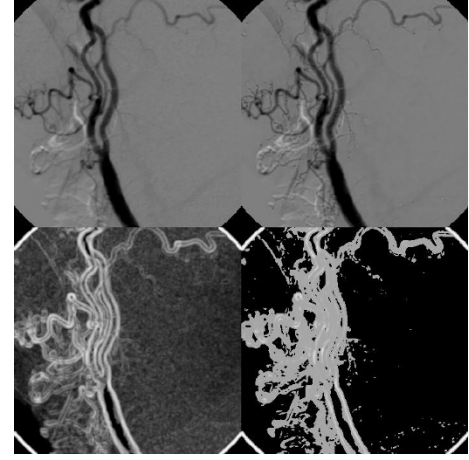

FIGURE 6: Residue-image processing applied to a subtractive angiography image. The top row shows the images, while the bottom row shows the corresponding residue amplitudes (A_r). The left and right columns correspond to the original and the processed image, respectively. The displayed residue amplitude images show $\log(1 + A_r)$ instead of A_r in order to better discrimination.

be approximated by a χ^2-distribution

$$P(E_r) = \frac{1}{E_o(q - 1)!} \left(\frac{E_r}{E_o} \right)^{q-1} \exp\left(-\frac{E_r}{E_o} \right), \qquad (30)$$

with $2q$ degrees of freedom. The corresponding PDF for the residue amplitude is

$$P(A_r) = \frac{2}{A_o(q - 1)!} \left(\frac{A_r}{A_o} \right)^{2q-1} \exp\left[-\left(\frac{A_r}{A_o} \right)^2 \right], \qquad (31)$$

where $E_o = A_o^2$. The average value and the variance in the energy PDF are equal to $\mu_e = q \cdot E_o$ and $\sigma_e^2 = q \cdot E_o^2$, respectively. The maximum values for the energy and amplitude PDFs occur at energy $E_m = E_o(q - 1)$ and amplitude
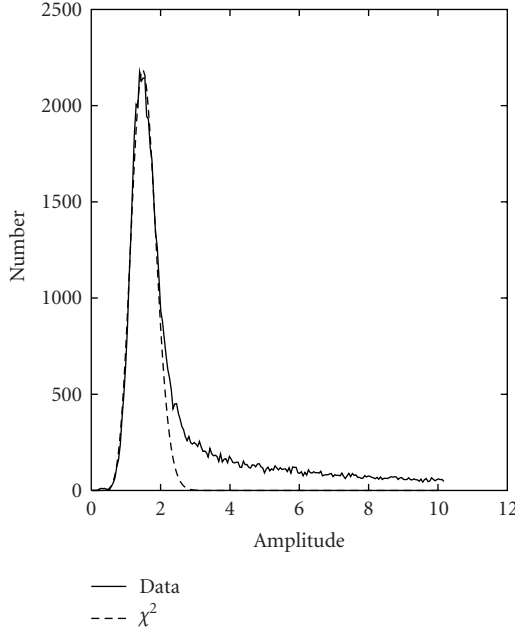
FIGURE 7: The drawn curve shows the histogram of the residue amplitude for the image and processing used in Figure 6. The dotted curve corresponds to optimal fitting of the χ^2-distribution ($q = 3.83$ and $A_o = 0.71$, resulting in a peak value at $A_m = 1.49$).

$A_m = A_o\sqrt{q - 0.5}$, respectively. Energy or amplitude histograms for an actual image can be compared against the above theoretical PDFs in order to find the PDF parameters E_o (or A_o) and q. Actual images of course also contain nonuniform regions. The corresponding responses for these image features will influence the shape of the energy histogram, especially for large energy (or amplitude) values. This implies that only the part of the histogram at low energy (or amplitude) values should be used for noise estimation and curve fitting against the theoretical PDF [24], as shown in Figure 7.

A parametrized PDF such as the χ^2-distribution can be used to perform K-means clustering on the residue amplitude or energy. For example, if the observed residue amplitude histogram $h(A_r)$ can be approximated by a sum of two components, that is,

$$h(A_r) = P_o \cdot P(A_r) + [h(A_r) - P_o \cdot P(A_r)], \qquad (32)$$

where $P(A_r)$ satisfies (31) and is associated with uniform regions, then maximum a posteriori classification will classify all positions $\mathbf{p}$ for which the residue amplitude satisfies $A_r(\mathbf{p}) < A_t$, with

$$P_o \cdot P(A_t) = h(A_t) - P_o \cdot P(A_t), \qquad (33)$$

as belonging to uniform regions. This provides us with an automated means of selecting the threshold A_t.

5. THE HERMITE TRANSFORMATION

5.1. Definition

We will now apply the theory of local signal decompositions to the specific case of a two-dimensional Gaussian window

$$w(x, y; \sigma) = \frac{1}{\sigma\sqrt{\pi}} \exp\left[-\frac{x^2 + y^2}{2\sigma^2} \right] \qquad (34)$$

with standard deviation equal to σ. Note that this window function is both separable, that is, can be written as the product of two Gaussian functions in x- and y-directions, and *circularly symmetric*, that is, its value only depends on the distance $r = \sqrt{x^2 + y^2}$ from the origin. Especially the circular symmetry will turn out to be essential for the discussion that follows.

We select the basis functions of the signal decomposition (which were denoted by φ_n in Section 2) to be orthonormal polynomials for the Gaussian window [2]. Because this window is separable, these orthonormal polynomials are also separable, that is,

$$p_{n-m,m}(x, y; \sigma) = g_{n-m}\left(\frac{x}{\sigma}\right) \cdot g_m\left(\frac{y}{\sigma}\right), \qquad (35)$$

with

$$g_n(x) = \frac{1}{\sqrt{2^n n!}} H_n(x), \qquad (36)$$

where the standard notation is used to denote the Hermite polynomial $H_n(x)$ of order n, for $n = 0, 1, \dots$. For the discussion in this paper, we will especially need the polynomials up to degree 2, that is,

$$g_0(x) = 1, \qquad g_1(x) = \sqrt{2}x, \qquad g_2(x) = \frac{2x^2 - 1}{\sqrt{2}}. \qquad (37)$$

The resulting image decomposition at the generic[1] lattice position $(0, 0)$ satisfies

$$w(x, y; \sigma)\left[l(x, y) - \sum_{n=0}^{\infty} \sum_{m=0}^{n} l_{n-m,m} \cdot p_{n-m,m}(x, y) \right] = 0, \qquad (38)$$

where equality means convergence in the mean-square sense.

The *Hermite coefficients* [$l_{n-m,m}$; $n = 0, \dots, \infty$; $m = 0, \dots, n$] are derived by filtering the spatial image representation $l(x, y)$ with *derivative-of-Gaussian* filters with impulse responses

$$d_{n-m,m}(x, y; \sigma)$$
$$= g_{n-m}\left(-\frac{x}{\sigma} \right) g_m\left(-\frac{y}{\sigma} \right) \cdot w^2(x, y; \sigma)$$
$$= \frac{1}{\sqrt{2^n (n-m)! m!}} \frac{\partial^n}{\partial (x/\sigma)^{n-m} \partial (y/\sigma)^m} w^2(x, y; \sigma), \qquad (39)$$

---

[1] In order to simplify the notation we discuss only the contribution at lattice position $(0, 0)$; the generalization to contributions $w(x-p, y-q; \sigma) \cdot l(x, y)$ at arbitrary lattice positions $(p, q) \in \mathcal{P}$ is straightforward.

or filter kernels $d_{n-m,m}(-x, -y; \sigma)$, and selecting the filtered outputs at the lattice positions. For the generic lattice position $(0, 0)$ this results in

$$l_{n-m,m} = \int_{-\infty}^{+\infty} dx \int_{-\infty}^{+\infty} dy \, l(x, y) \cdot d_{n-m,m}(-x, -y; \sigma). \quad (40)$$

The fact that the filter functions in (39) can be expressed as derivatives of Gaussians is a consequence of a well-known property of Hermite polynomials, that is,

$$e^{-x^2} H_n(x) = (-1)^n \frac{d^n}{dx^n} e^{-x^2}. \quad (41)$$

This property implies that the Hermite coefficients themselves can also be interpreted as partial derivatives

$$l_{n-m,m} = \frac{1}{\sqrt{2^n (n-m)! m!}} \left[\frac{\partial^n}{\partial (x/\sigma)^{n-m} \partial (y/\sigma)^m} l(x, y; \sigma) \right]_{x=y=0} \quad (42)$$

to the smoothed image

$$l(x, y; \sigma) = l(x, y) \star w^2(x, y; \sigma) \quad (43)$$

that is obtained by applying a Gaussian filter $w^2(x, y; \sigma)$ with standard deviation equal to $\sigma/\sqrt{2}$ to the original image $l(x, y)$.

Note that the contribution at the generic lattice position $(0, 0)$ to the overall synthesized image is

$$\begin{aligned} l_w(x, y) &= w^2(x, y; \sigma) \cdot l(x, y) \\ &= \sum_{n=0}^{\infty} \sum_{m=0}^{n} l_{n-m,m} \cdot d_{n-m,m}(-x, -y; \sigma), \end{aligned} \quad (44)$$

while the local energy at this position is equal to

$$E = \iint w^2(x, y) \cdot l^2(x, y) dx \, dy = \sum_{n=0}^{\infty} \sum_{m=0}^{n} l_{n-m,m}^2. \quad (45)$$

The residue energy $E_r = E - l_{00}^2$ or residue amplitude $A_r = \sqrt{E_r}$ can again be used to differentiate between uniform and nonuniform regions in the image.

In Figure 8 we show an artificial image containing several interesting features such as edges, lines, corners, T-junctions, X-junctions, circles, and noise. We will use this image throughout this paper to illustrate different alternative forms of the Hermite transformation because it is much easier to judge the properties of the Hermite transformation on this image than on a natural image.

In Figure 9 we show the residue amplitude and the Hermite coefficients up to order 2 for a Gaussian window with spread $\sigma = 2.5$ times the sampling distance. The coefficients $l_{n-m,m}$ for all positions are grouped into subimages. The (fuzzy) derivatives of order $n - m$ along the x-direction and order m along the y-direction are obtained by filtering the original image with the derivative-of-Gaussian filter $d_{n-m,m}(x, y)$ and (optionally) subsampling the filtered image. No subsampling has been applied in the case of Figure 9. One important advantage of the derivative-of-Gaussian filters is that they are separable. This means that they can be implemented very efficiently by first filtering along the x-direction and subsequently along the y-direction (or vice versa).
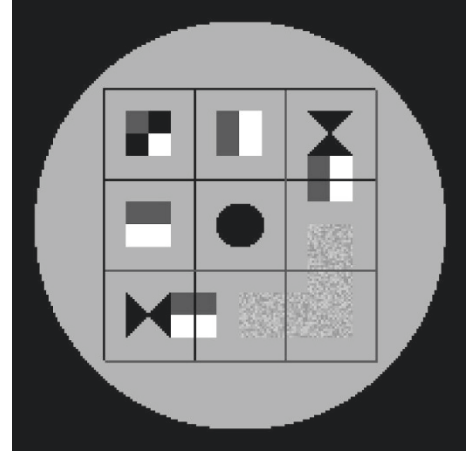

FIGURE 8: Original image used in the Hermite analyses.

5.2. Scale space

Derivative-of-Gaussian filters play a key role in the theory of scale space [32, 37, 38]. Scale space represents a systematic way of studying structure in images. An image is regarded as a surface that can be analyzed at varying degrees of detail. To reduce the level of detail, the image is first passed through a Gaussian filter. The smoothed surface is then described using traditional methods from differential geometry (see also Section 7), a well-studied discipline of mathematics [39]. Many interesting surface properties can be expressed as combinations of the local partial derivatives $l_{n-m,m}$ of (42) to the smoothed image $l(x, y; \sigma)$. It is well known that this smoothed image is the solution to a diffusion equation

$$\frac{\partial}{\partial t} u(x, y, t) = \left[\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \right] u(x, y, t), \quad (46)$$

with initial condition

$$u(x, y, 0) = l(x, y) \quad (47)$$

that is stopped after time $t = T = (\sigma/2)^2$. Fast recursive methods for numerically solving this diffusion equation [40] may be used to determine $u(x, y, T) = l(x, y; 2\sqrt{T})$ for large values of T. In the Hermite transformation the operations of smoothing and differentiating are combined, so that the Hermite transformation corresponds to an analysis at a specified level in scale space.

Very often, the most efficient way for calculating the Hermite coefficients is by combining diffusion and filtering, as shown in Figure 10. Indeed, if the input image $l(x, y)$ is first diffused up to time $T = (\sigma_D/2)^2$, and subsequently analyzed by a Hermite transformation with parameter σ_H, then the resulting coefficients at the origin, that is,

$$u_{n-m,m} = [d_{n-m,m}(x, y; \sigma_H) \star u(x, y; T)]_{x=y=0} = \left(\frac{\sigma_H}{\sigma} \right)^n l_{n-m,m}, \quad (48)$$
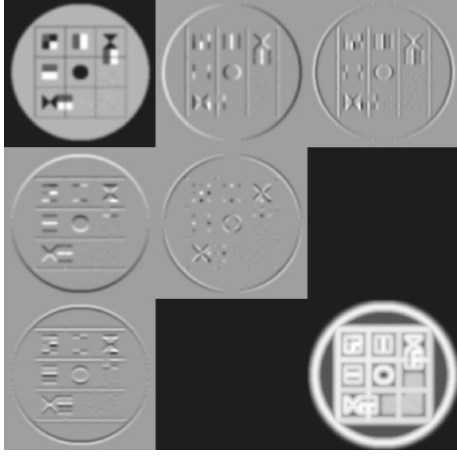
FIGURE 9: Hermite coefficients $l_{n-m,m}$, for $m = 0, \ldots, n$, up to order $n = 2$ for a Gaussian window with spread $\sigma = 2.5$ times the sampling distance. The image in the lower-right corner is the residue amplitude A_r. To relate the mathematical results in this paper to the figures, it should be realized that the x-axis points to the right and the y-axis downwards. Moreover, the coefficients of order n, which are located in the nth diagonal, are amplified by the same factor β_n for display purposes. The factors β_n are chosen independently for each order n, in such a way that the complete display range is used.
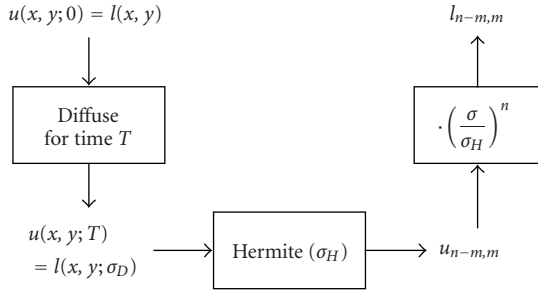

FIGURE 10: The Hermite coefficients for a large standard deviation equal to σ can be obtained by a three-step process. First, a diffusion is applied to the original image $u(x, y; 0) = l(x, y)$ for a time T, resulting in the diffused image $u(x, y; T) = l(x, y; \sigma_D)$, where $\sigma_D = 2\sqrt{T}$. Second, a Hermite transformation with a small standard deviation equal to σ_H, where $\sigma_H^2 = \sigma^2 - \sigma_D^2$, is applied to the diffused image, resulting in the coefficients $u_{n-m,m}$, for $m = 0, \ldots, n$ and $n = 0, 1, 2 \ldots$. Third, these coefficients of order n are rescaled by a factor $(\sigma/\sigma_H)^n$ to obtain the required Hermite coefficients $l_{n-m,m}$.

belong to a Hermite transformation with parameter

$$\sigma = \sqrt{\sigma_D^2 + \sigma_H^2} = \sqrt{4T + \sigma_H^2}. \qquad (49)$$

An upper limit on σ_H allows to limit the size of the derivatives-of-Gaussian filters, since this filter size is proportional to σ_H. The number of computations required in the filtering stage increases (linearly) with this filter size. A lower limit on σ_H (typically, σ_H larger than 1.5 times the sampling distance [13]) is required to guarantee that the derivative operations can be closely approximated for digital images. Most of the time we use a well-designed set of

derivatives-of-Gaussian filters (for σ_H equal to 1.8 times the sampling distance), and vary the diffusion time T to accomplish the required value of σ.

5.3. Angular functions

An important step towards obtaining a better insight into the Hermite transformation is to express the filtering that is performed on the input image in the Fourier domain. The Fourier transformations of the filter functions, expressed in polar coordinates, are

$$\hat{D}_{n-m,m}(\omega, \phi; \sigma) = \hat{D}_n(\omega; \sigma) \cdot \alpha_{n-m,m}(\phi), \qquad (50)$$

where

$$\hat{D}_n(\omega; \sigma) = \frac{1}{\sqrt{2^n n!}} (-j\omega\sigma)^n \exp\left[-\frac{(\omega\sigma)^2}{4} \right] \qquad (51)$$

is the 1D Fourier transformation of the nth-order Gaussian derivative

$$D_n(r; \sigma) = \frac{1}{\sqrt{2^n n!}} \frac{d^n}{d(r/\sigma)^n} \left[\frac{1}{\sigma\sqrt{\pi}} \exp\left(-\frac{r^2}{\sigma^2} \right) \right], \qquad (52)$$

and

$$\alpha_{n-m,m}(\phi) = \sqrt{\frac{n!}{(n-m)!m!}} \cos^{n-m}\phi \cdot \sin^m\phi, \qquad (53)$$

for $m = 0, \ldots, n$, are the *Cartesian angular functions* of order n [2, 20]. The shape of these Cartesian angular functions can be judged very well in Figure 9 from the angular modulation of the coefficients along the edge of the circle image. The Cartesian angular functions of order $n = 1$ are simply $\alpha_{10}(\phi) = \cos\phi$ and $\alpha_{01}(\phi) = \sin\phi$, while the Cartesian angular functions of order $n = 2$, that is, $\alpha_{20}(\phi) = \cos^2\phi$, $\alpha_{11}(\phi) = \sqrt{2}\cos\phi\sin\phi$, and $\alpha_{02}(\phi) = \sin^2\phi$, are plotted in Figure 11.

The Hermite coefficients can be derived from the Fourier image representation $\hat{l}(\omega_x, \omega_y)$ through

$$l_{n-m,m} = \int_{-\infty}^{+\infty} \frac{d\omega_x}{2\pi} \int_{-\infty}^{+\infty} \frac{d\omega_y}{2\pi} \hat{l}(\omega_x, \omega_y) \cdot \hat{d}_{n-m,m}(\omega_x, \omega_y; \sigma) \qquad (54)$$

or, alternatively, by changing to polar coordinates, through

$$l_{n-m,m} = \frac{1}{2\pi} \int_0^{2\pi} d\phi \hat{l}_n(\phi) \cdot \alpha_{n-m,m}(\phi) \qquad (55)$$

for $m = 0, \ldots, n$, where

$$\hat{l}_n(\phi) = \frac{1}{2\pi} \int_0^\infty \omega \, d\omega \, \hat{l}(\omega, \phi) \cdot \hat{D}_n(\omega; \sigma) \qquad (56)$$

is called the "*image*" *angular function* of order n [20]. This function is real if the original image $l(x, y)$ is real. It moreover satisfies $\hat{l}_n(\phi + \pi) = (-1)^n \cdot \hat{l}_n(\phi)$, so that it is odd or even symmetric, depending on whether the order n is odd or even, respectively. Obviously, the Hermite coefficients $[l_{n-m,m}, m = 0, \ldots, n]$ of the same order n can all be derived

- —— (2, 0)
- - - - (1, 1)
- ······ (0, 2)
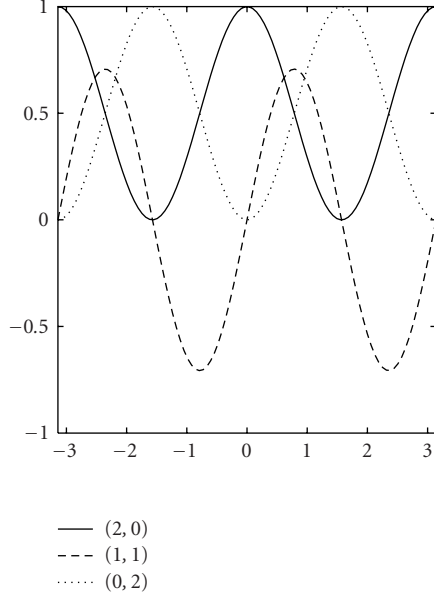
FIGURE 11: Cartesian angular functions of order $n = 2$.

from the angular function $\widehat{l}_n(\phi)$. One advantage of defining the "image" angular function as an intermediate step in the analysis is that it becomes straightforward to extend many of the results below, that is, those that only depend on the properties of the angular components of the filters, to filter banks in which the radial components of the filters, that is, $\widehat{D}_n(\omega; \sigma)$, are not derivatives-of-Gaussians.

The Fourier transformation of the local image $l_w(x, y) = w^2(x, y; \sigma) \cdot l(x, y)$ can be expressed as

$$\widehat{l}_w(\omega, \phi) = \widehat{l}(\omega, \phi) \star \widehat{D}_0(\omega) = \sum_{n=0}^{\infty} \widehat{D}_n^{\star}(\omega) \cdot l_n(\phi), \qquad (57)$$

where

$$l_n(\phi) = \sum_{m=0}^{n} l_{n-m,m} \cdot \alpha_{n-m,m}(\phi) \qquad (58)$$

is called the "*derivative*" *angular function* of order n. Since the interest here will be mainly in the derivative angular functions of order $n = 1$ and $n = 2$, the following explicit expressions

$$l_1(\phi) = l_{10} \cdot \cos \phi + l_{01} \cdot \sin \phi,$$
$$l_2(\phi) = l_{20} \cdot \cos^2 \phi + l_{11} \cdot \sqrt{2} \cos \phi \sin \phi + l_{02} \cdot \sin^2 \phi \qquad (59)$$

are provided.

5.4. One-dimensional images

The Hermite coefficients at the generic lattice position $(0, 0)$ describe the local image $l_w(x, y) = w^2(x, y; \sigma) \cdot l(x, y)$. An important specific case of local patterns in 2D are 1D patterns, that is, patterns that vary only in one direction (and

are constant along the orthogonal direction). In natural images, many of the image details that are of prime importance, such as edges and lines, can be locally described as one-dimensional patterns, at least within the analysis window $w(x, y; \sigma)$. Since analysis, coding, and processing of such patterns are considered important, it is worthwhile to investigate them separately.

The angular function for a locally one-dimensional (1D) image with normal along direction $(\cos \phi_o, \sin \phi_o)$, that is, $l(x, y) = l(x \cos \phi_o + y \sin \phi_o)$, has only contributions in the directions ϕ_o and $\phi_o + \pi$ and is therefore of the form

$$\widehat{l}_n(\phi) = \pi l_n [\delta(\phi - \phi_o) + (-1)^n \delta(\phi - (\phi_o + \pi))], \qquad (60)$$

where l_n is the nth-order Hermite coefficient of the 1D signal $l(r)$.

For a 1D image we can express the Hermite coefficients as

$$l_{n-m,m} = l_n \cdot \alpha_{n-m,m}(\phi_o), \qquad (61)$$

so that the "derivative" angular function becomes

$$l_n(\phi) = l_n \cdot \sum_{m=0}^{n} \alpha_{n-m,m}(\phi_o) \cdot \alpha_{n-m,m}(\phi) = l_n \cdot \cos^n(\phi - \phi_o). \qquad (62)$$

6. LOCAL COORDINATE AXES

6.1. Effect on Hermite coefficients

In scale-space theory [32], geometric features of the smoothed surface $l(x, y; \sigma)$ as a function of the position (x, y) (and scale σ) are used for image analysis [41, 42]. The well-established mathematical discipline of differential geometry [43] is especially concerned with how such geometric features can be derived, and some of the major results of applying differential geometry for image feature extraction will be discussed in Section 7.

Adopting local coordinate axes

$$\begin{bmatrix} \widetilde{x} \\ \widetilde{y} \end{bmatrix} = \begin{bmatrix} \cos \phi_o & \sin \phi_o \\ -\sin \phi_o & \cos \phi_o \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix} \qquad (63)$$

is one of the basic operations in differential geometry because many expressions for important surface properties can be simplified greatly in an adaptive coordinate system. For example, for the one-dimensional patterns introduced above, we will see shortly that aligning the $\widetilde{x}$-axis with the direction of the pattern has distinct advantages.

We analyze the behavior of the angular functions $l_n(\phi)$ and $\widehat{l}_n(\phi)$ under rotation of the coordinate axes. If we rotate the coordinate axes clockwise (or the image anticlockwise) through an angle ϕ_o (in the left-handed coordinate system that is used in the figures), that is, if we change $\widehat{l}_n(\phi)$ to $\widehat{l}_n(\phi + \phi_o)$, then the Hermite coefficients in the local coordinate system $(\widetilde{x}, \widetilde{y})$ are equal to

$$l_{n-m,m}(\phi_o) = \frac{1}{2\pi} \int_0^{2\pi} d\phi \widehat{l}_n(\phi) \cdot \alpha_{n-m,m}(\phi - \phi_o), \qquad (64)$$

for $m = 0, \ldots, n$, and are related to the coefficients $[l_{n-m,m}, m = 0, \ldots, n]$ in the global coordinate system (x, y) by a linear transformation with parameter ϕ_o. The linear nature of the transformation is a consequence of the fact that both $\cos(\phi - \phi_o)$ and $\sin(\phi - \phi_o)$ can be expressed as linear combinations of $\cos \phi$ and $\sin \phi$. Specifically, since

$$\alpha_{n,0}(\phi - \phi_o) = \cos^n(\phi - \phi_o) = \sum_{m=0}^{n} \alpha_{n-m,m}(\phi_o) \cdot \alpha_{n-m,m}(\phi), \tag{65}$$

we obtain that the derivative of order n along direction ϕ_o is given by $l_{n,0}(\phi_o) = l_n(\phi_o)$, which explains the name "derivative" angular function for $l_n(\phi)$. Based on the $n + 1$ partial derivatives $[l_{n-m,m}; m = 0, \ldots, n]$, we can hence derive the directional derivative $l_n(\phi_o)$ of order n for any direction ϕ_o [20, 44]. The derivative angular function of order n can of course also be expressed in terms of the Hermite coefficients in the rotated coordinate frame, that is,

$$l_n(\phi) = \sum_{m=0}^{n} l_{n-m,m}(\phi_o) \cdot \alpha_{n-m,m}(\phi - \phi_o). \tag{66}$$

Note that the linear transformation of the Hermite coefficients under rotation of the coordinate system requires that the image be locally approximated by a polynomial within a circularly symmetric window $w(x, y; \sigma)$. Rotating the coordinate axes can have a complicated effect on the Hermite coefficients, especially if the order n is high. A systematic analysis for arbitrary orders has been performed [20], but is beyond the scope of this paper. In the remainder, we will only need the results for orders $n = 1$, that is,

$$\begin{bmatrix} l_{10}(\phi_o) \\ l_{01}(\phi_o) \end{bmatrix} = \begin{bmatrix} \cos \phi_o & \sin \phi_o \\ -\sin \phi_o & \cos \phi_o \end{bmatrix} \cdot \begin{bmatrix} l_{10} \\ l_{01} \end{bmatrix}, \tag{67}$$

and $n = 2$, that is,

$$\begin{bmatrix} l_{20}(\phi_o) \\ l_{11}(\phi_o) \\ l_{02}(\phi_o) \end{bmatrix} = T_2^t \cdot \begin{bmatrix} \cos(2\phi_o) & 0 & \sin(2\phi_o) \\ 0 & 1 & 0 \\ -\sin(2\phi_o) & 0 & \cos(2\phi_o) \end{bmatrix} \cdot T_2 \cdot \begin{bmatrix} l_{20} \\ l_{11} \\ l_{02} \end{bmatrix}, \tag{68}$$

where

$$T_2 = \begin{bmatrix} \dfrac{1}{\sqrt{2}} & 0 & -\dfrac{1}{\sqrt{2}} \\ \dfrac{1}{\sqrt{2}} & 0 & \dfrac{1}{\sqrt{2}} \\ 0 & 1 & 0 \end{bmatrix} \tag{69}$$

is a unitary transformation matrix (with determinant equal to -1). These results for orders $n = 1, 2$ can also be obtained by straightforward manipulation of the goniometric

functions involved in (64), for example,

$$\begin{aligned} l_{1,0}(\phi_o) &= \frac{1}{2\pi} \int_0^{2\pi} d\phi \hat{l}_n(\phi) \cos(\phi - \phi_o) \\ &= \frac{1}{2\pi} \int_0^{2\pi} d\phi \hat{l}_n(\phi) \cos \phi \cdot \cos \phi_o \\ &\quad + \frac{1}{2\pi} \int_0^{2\pi} d\phi \hat{l}_n(\phi) \sin \phi \cdot \sin \phi_o \\ &= l_{10} \cdot \cos \phi_o + l_{01} \cdot \sin \phi_o. \end{aligned} \tag{70}$$

Since rotation of an image within a circularly symmetric window does not alter the signal energy contained in this window, we obtain that the local energy up to order N satisfies

$$E_N = \sum_{n=0}^{N} \sum_{m=0}^{n} l_{n-m,m}^2 = \sum_{n=0}^{N} \sum_{m=0}^{n} l_{n-m,m}^2(\phi_o), \tag{71}$$

for all $N \geq 0$. This also explains why the above transformation matrices are unitary matrices.

6.2. Optimizing energy compaction

We now discuss how a suitable orientation for the local coordinate axes can be derived from the Hermite coefficients. For a 1D image with orientation ϕ_o, we obtain that

$$l_{n-m,m} = l_n \cdot \alpha_{n-m,m}(\phi_o) \quad \text{or} \quad l_{n-m,m}(\phi_o) = l_n \cdot \alpha_{n-m,m}(0), \tag{72}$$

so that $l_{n,0}(\phi_o) = l_n(\phi_o) = l_n$ and $l_{n-m,m}(\phi_o) = 0$, for $m = 1, \ldots, n$, in this case. We can hence split the local energy up to order N as follows:

$$\begin{aligned} E_N &= l_{00}^2 + E_{N,1d}(\phi_o) + E_{N,2d}(\phi_o) \\ &= l_{00}^2 + \sum_{n=1}^{N} l_{n,0}^2(\phi_o) + \sum_{n=1}^{N} \sum_{m=1}^{n} l_{n-m,m}^2(\phi_o), \end{aligned} \tag{73}$$

where $E_{N,1d}(\phi_o)$ and $E_{N,2d}(\phi_o)$ are the local 1D and 2D energies up to order N along direction ϕ_o, respectively. We therefore propose to use a strategy in which ϕ_o is selected such that $E_{N,1d}(\phi_o)$ is maximized.

In Figure 12, we show an adaptive Hermite transformation of order $N = 2$ in which the local orientation ϕ_o has been selected by optimizing $E_{N,1d}(\phi_o)$. Note the perfect energy compaction that is obtained for 1D patterns. The 1D-amplitude $A_{N,1d}(\phi_o) = [E_{N,1d}(\phi_o)]^{1/2}$ and 2D-amplitude $A_{N,2d}(\phi_o) = [E_{N,2d}(\phi_o)]^{1/2}$ are very useful features, since they can be used to distinguish between 1D and 2D features in an image. The 1D features that are detected in this way can be classified based on the set of 1D Hermite coefficients $l_{n,0}(\phi_o) = l_n$, for $n = 0, 1, \ldots$. This energy compaction and 1D pattern classification has been exploited in image coding [5, 6] and image-feature extraction [16, 18, 24]. The extracted features have in turn proven useful for single-ended quality predictions (i.e., quality predictions that do not require the original image) on blurred images [19] and JPEG-coded images [45].

6.3. Gradient direction

In differential geometry, the reference angle ϕ_o is often chosen such that the value of the derivative angular function of order $n = 1$, that is,

$$l_1(\phi_o) = l_{10} \cdot \cos\phi_o + l_{01} \cdot \sin\phi_o, \qquad (74)$$

is maximized. This is equivalent to optimizing the 1D energy up to order $N = 1$. The above angular function can be rewritten as

$$l_1(\phi_o) = l_1 \cdot \cos(\phi_o - \phi_d), \qquad (75)$$

with $l_1 = \sqrt{l_{10}^2 + l_{01}^2}$, provided that the angle ϕ_d satisfies

$$\tan\phi_d = \frac{l_{01}}{l_{10}}, \qquad (76)$$

or equivalently, $l_{10} = l_1 \cdot \cos\phi_d$ and $l_{01} = l_1 \cdot \sin\phi_d$. The maximum value obviously occurs for $\phi_o = \phi_d$. This direction that makes an angle ϕ_d with the horizontal axis is called the *gradient direction.*

In the neighbourhood of the generic lattice position $(0,0)$, the smoothed image can be expressed in both the global coordinate system (x, y) and the local coordinate system

$$\begin{bmatrix} x_d \\ y_d \end{bmatrix} = \begin{bmatrix} \cos\phi_d & \sin\phi_d \\ -\sin\phi_d & \cos\phi_d \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix}, \qquad (77)$$

that is, $l(x, y; \sigma) = l_d(x_d, y_d; \sigma)$ are two equivalent descriptions of the same local image. The derivatives for a Cartesian coordinate system that is aligned with this gradient direction can hence be expressed as

$$
\begin{aligned}
l_{n-m,m}^d &= \left. \frac{\partial^n}{\partial x_d^{n-m}\partial y_d^m} l_d(x_d, y_d; \sigma) \right|_{x_d=y_d=0} \\
&= \frac{\sqrt{2^n (n-m)! m!}}{\sigma^n} l_{n-m,m}(\phi_d),
\end{aligned}
\qquad (78)
$$

for $n = 0, 1, 2, \ldots$ and $m = 0, \ldots, n$. Selecting the reference angle along the gradient direction guarantees that $l_{01}^d = 0$. The value

$$l_{10}^d = \frac{\sqrt{2}}{\sigma} l_1 = \frac{\sqrt{2}}{\sigma} \sqrt{l_{10}^2 + l_{01}^2} \qquad (79)$$

is called the *gradient amplitude.* The following explicit relationships

$$l_{20}^d = \frac{2\sqrt{2}}{\sigma^2} l_{20}(\phi_d), \qquad l_{11}^d = \frac{2}{\sigma^2} l_{11}(\phi_d), \qquad l_{02}^d = \frac{2\sqrt{2}}{\sigma^2} l_{02}(\phi_d) \qquad (80)$$

exist between the Hermite coefficients of order 2 and the second-order partial derivatives to the smoothed image $l_d(x_d, y_d; \sigma)$ for coordinate axes (x_d, y_d) oriented along gradient direction.
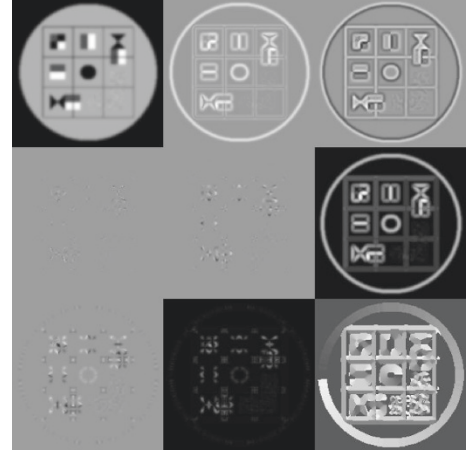

FIGURE 12: Hermite coefficients $l_{n-m,m}(\phi_o)$, for $m = 0, \ldots, n$, up to order $n = 2$ for adaptive coordinate axes and a Gaussian window with spread $\sigma = 2.5$ times the sampling distance. The image in the lower-right corner codes the local orientation ϕ_o. The image at the right in the middle row is the 1D-amplitude $A_{2,1d}(\phi_o)$, while the image at the bottom in the middle column is the 2D-amplitude $A_{2,2d}(\phi_o)$.

7. GEOMETRIC FEATURES

7.1. Planar curves

The geometric features that we will derive from the Hermite coefficients will all be characteristics of curves that are in some way defined by the surface function $l(x, y; \sigma)$. We will only consider curves that belong to a 2D plane, and therefore start by summarizing some basic properties of such *planar curves.* We denote by $\mathbf{c}(t) = (u(t), v(t))$ the coordinates traced out by a curve segment in an orthogonal coordinate system (u, v) of a 2D plane (this plane need not be the image plane (x, y)). The parameter t is defined for an interval of the real axis that includes 0, while $\mathbf{c}(0) = (u(0), v(0))$ is the point at which we want to measure the curve characteristics. We assume that the functions $u(t)$ and $v(t)$ that relate the parameter t to the curve coordinates are at least twice differentiable.

We will use two well-known properties of planar curves that are both illustrated in Figure 13. First, the tangent line to the curve at position $\mathbf{c}(t)$ has the direction indicated by the *tangent vector*

$$\mathbf{c}_t(t) = \frac{1}{\left([u'(t)]^2 + [v'(t)]^2 \right)^{1/2}} (u'(t), v'(t)). \qquad (81)$$

This vector can be used as the first vector in a right-handed orthonormal coordinate system. The second coordinate vector thus defined is called the *normal vector* and can be expressed as follows:

$$\mathbf{c}_n(t) = \frac{1}{\left([u'(t)]^2 + [v'(t)]^2 \right)^{1/2}} (-v'(t), u'(t)). \qquad (82)$$

As we move along the curve, the direction of the tangent (and normal) vector will change. The next theorem on planar
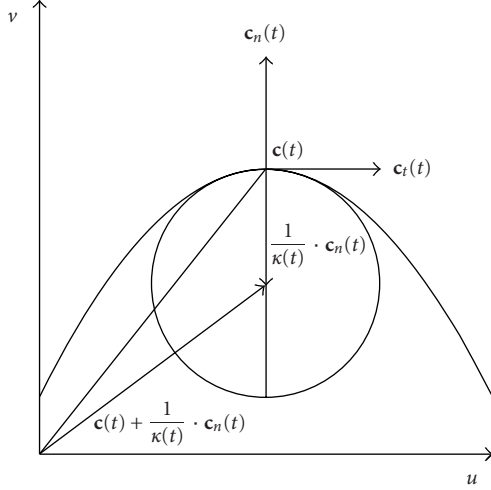
FIGURE 13: The tangent vector $\mathbf{c}_t(t)$ at the curve $\mathbf{c}(t)$ can be combined with the normal vector $\mathbf{c}_n(t)$ to obtain a right-handed orthonormal coordinate system. The circle with radius $1/|\kappa(t)|$ has second-order contact with the curve $\mathbf{c}(t)$.

curves tells us how to measure this rate of change. More specifically, there exists a circle that is tangent to the curve at $\mathbf{c}(t)$ and for which the normal vector to the curve turns at the same rate as the normal vector to the circle when moving in the direction of the tangent vector (i.e., this circle has second-order contact with the curve). This circle has radius $1/|\kappa(t)|$ and its center is located at position

$$\mathbf{c}(t) + \frac{1}{\kappa(t)}\mathbf{c}_n(t), \tag{83}$$

where

$$\kappa(t) = \frac{u'(t) \cdot v''(t) - u''(t) \cdot v'(t)}{\left(\left[u'(t)\right]^2 + \left[v'(t)\right]^2\right)^{3/2}} \tag{84}$$

is called the *curvature* at position $\mathbf{c}(t)$ [43]. A positive or negative curvature implies that the curve bends towards or away from the normal vector, respectively.

7.2. Isophote curvature

A first planar curve that can be derived from the surface $z = l_d(x_d, y_d; \sigma)$ is the *isophote* through the origin, defined by

$$l_d(u(t), v(t); \sigma) = l_d(u(0), v(0); \sigma) \tag{85}$$

for some parameter interval including $t = 0$. This is the curve that arises if the image surface is intersected with a horizontal plane through the surface at the point of interest. The image value along the isophote does not change as a function of t,

so that

$$\left.\frac{d}{dt} l_d(u(t), v(t); \sigma)\right|_{t=0} = l_{10}^d \cdot u' + l_{01}^d \cdot v',$$

$$\left.\frac{d^2}{dt^2} l_d(u(t), v(t); \sigma)\right|_{t=0}$$
$$= l_{20}^d \cdot (u')^2 + 2l_{11}^d \cdot u'v' + l_{02}^d \cdot (v')^2 + l_{10}^d \cdot u'' + l_{01}^d \cdot v'' \tag{86}$$

are both equal to zero. Since the coordinate system is aligned along the gradient direction, that is, $l_{01}^d = 0$, we obtain that $u'(0) = 0$ and $u''(0) = -[v'(0)]^2 \cdot l_{02}^d/l_{10}^d$, which in turn leads to the expression

$$\kappa_i = -\frac{u''(0)}{[v'(0)]^2} = \frac{l_{02}^d}{l_{10}^d} = \frac{2}{\sigma}\frac{l_{02}(\phi_d)}{l_{10}(\phi_d)}, \tag{87}$$

for the *isophote curvature* at the origin.

Since isophotes will have a high curvature near a corner, such as the one shown in Figure 14, the isophote curvature can assist in corner detection. More specifically, finding local maxima in the derived *isophote curvature measure*

$$m_i = \left| \kappa_i \cdot (l_{10}^d)^3 \right|^{1/3} = \left| l_{02}^d \cdot (l_{10}^d)^2 \right|^{1/3}, \tag{88}$$

which also contains contrast information by including the image gradient l_{10}^d, is a frequently used method for selecting corner candidates. Figure 14 illustrates that, although the position of a maximum in the isophote curvature measure is usually very close to the true corner position, both positions will usually not coincide. This implies that a maximum in the isophote curvature measure can at best be used as an indication for the presence of a corner, and that a more in-depth analysis is required, for instance to update the position estimate of the corner feature [32].

The isophote curvature measure for the entire test image of Figure 8 is shown in Figure 15(a).

7.3. Flowline curvature

A second planar curve that can be derived from the surface $z = l_d(x_d, y_d; \sigma)$ is the *flowline* through the origin, defined by the demand that the tangent to the flowline is equal to the image gradient, that is,

$$u'(t) = \left.\frac{\partial}{\partial x_d} l_d(x_d, y_d; \sigma)\right|_{x_d=u(t), y_d=v(t)},$$
$$v'(t) = \left.\frac{\partial}{\partial y_d} l_d(x_d, y_d; \sigma)\right|_{x_d=u(t), y_d=v(t)}, \tag{89}$$

for some parameter interval including $t = 0$. Note that, at the origin, the flowline is orthogonal to the isophote.

Evaluating the second derivatives to the flowline at the origin results in

$$u''(0) = l_{20}^d \cdot u'(0) + l_{11}^d \cdot v'(0) = l_{20}^d \cdot l_{10}^d,$$
$$v''(0) = l_{11}^d \cdot u'(0) + l_{02}^d \cdot v'(0) = l_{11}^d \cdot l_{10}^d \tag{90}$$
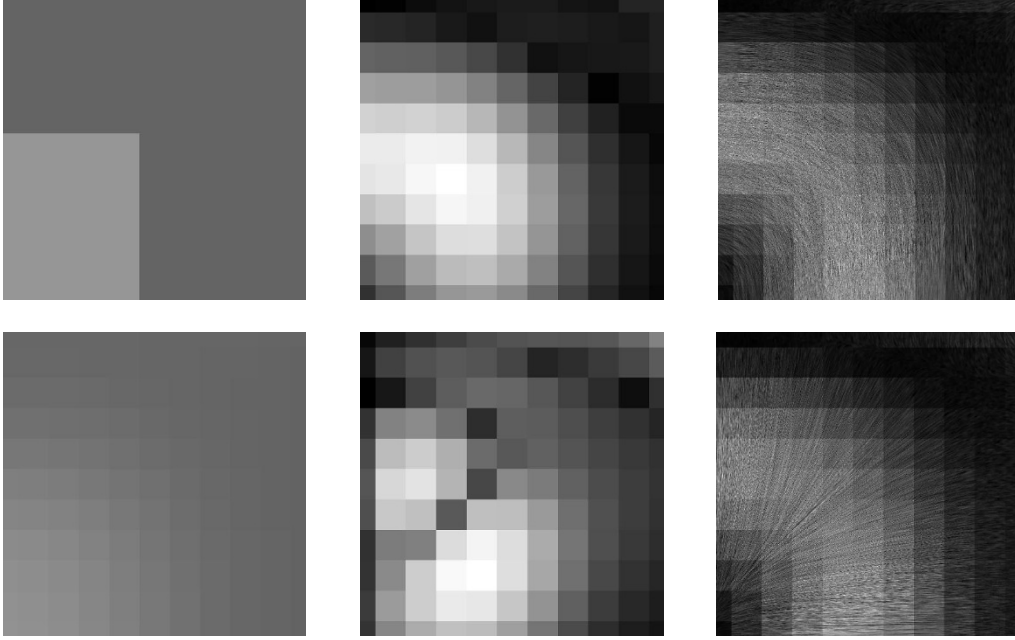
FIGURE 14: Curvature measures for a corner feature extracted from **Figure 8**. Upper left: original image, lower left: smoothed image, upper middle: isophote curvature measure $|m_i|$, and lower middle: flowline curvature measure $|m_f|$. The rightmost figures use the "spot noise" technique [46] to visualize the isophote vector field (upper) and the flowline (or gradient) vector field that is orthogonal to it (lower).

which immediately leads to the expression

$$\kappa_f = \frac{v''(0)}{\left[u'(0)\right]^2} = \frac{l_{11}^d}{l_{10}^d} = \frac{\sqrt{2}}{\sigma} \frac{l_{11}(\phi_d)}{l_{10}(\phi_d)} \qquad (91)$$

for the *flowline curvature* at the origin [32], since $u'(0) = l_{10}^d$ and $v'(0) = l_{01}^d = 0$.

The derived *flowline curvature measure*

$$m_f = \left| \kappa_f \cdot \left(l_{10}^d\right)^3 \right|^{1/3} = \left| l_{11}^d \cdot \left(l_{10}^d\right)^2 \right|^{1/3} \qquad (92)$$

that also depends on the image gradient can be shown to react especially to junctions, such as the T-junction shown in Figure 16. Again note that the position of the maximum in the flowline curvature measure does not coincide exactly with the true position of the T-junction. The flowline curvature measure for the entire test image of Figure 8 is shown in Figure 15(b).

In conclusion, the feature vector $(|m_i|, |m_f|)$ is very useful for the (inital) detection and classification of corners and junctions. Candidate corner/junction positions are obtained by finding local maxima (and minima) in these curvature measures. The next section discusses some geometric image features that can be used to find such local extrema in (feature) images.

7.4. Intersection and surface curvatures

7.4.1. Intersection curvatures

A third kind of planar curve that can be derived from the surface $z = l(x, y; \sigma)$ is the intersection curve with a vertical plane through the origin in the direction that makes an angle ϕ with the global x-axis (or an angle $\phi - \phi_d$ with the local gradient axis x_d). This curve can be parametrized by $u(t) = t$ and

$$\begin{aligned} v(t) &= l(t \cdot \cos \phi, t \cdot \sin \phi; \sigma) \\ &= l_d\big(t \cdot \cos(\phi - \phi_d), t \cdot \sin(\phi - \phi_d); \sigma\big). \end{aligned} \qquad (93)$$

The first-order derivatives for $t = 0$ are equal to $u'(0) = 1$ and

$$v'(0) = l_1(\phi) \frac{\sqrt{2}}{\sigma}, \qquad (94)$$

while higher-order derivatives satisfy $u^{(n)}(0) = 0$ and

$$v^{(n)}(0) = l_n(\phi) \frac{\sqrt{2^n n!}}{\sigma^n}, \qquad (95)$$

where $l_n(\phi)$ is the "derivative" angular function of order n. The intersection curve hence has an *intersection curvature* in the direction ϕ equal to

$$\kappa(\phi) = \frac{\sigma l_2(\phi)}{\left[(\sigma/\sqrt{2})^2 + (l_1(\phi))^2\right]^{3/2}}. \qquad (96)$$

The intersection curvature will vary between $\kappa_{\min} = \kappa(\phi_{\min})$ and $\kappa_{\max} = \kappa(\phi_{\max})$. These minimum and maximum curvatures can be used directly as image features or they can be combined into $\kappa_l = \kappa_{\max} + \kappa_{\min}$, $\kappa_g = \kappa_{\max} \cdot \kappa_{\min}$, or $\kappa_r = |\kappa_{\max}^2 - \kappa_{\min}^2|$. These latter curvature measures are illustrated in the left part of Figure 17. An interpretation of
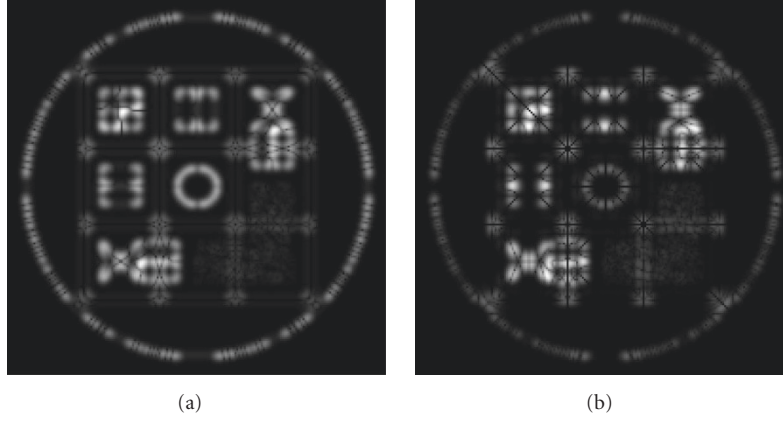
FIGURE 15: Curvature measures: (a) isophote curvature measure $|m_i|$ and (b) flowline curvature measure $|m_f|$.
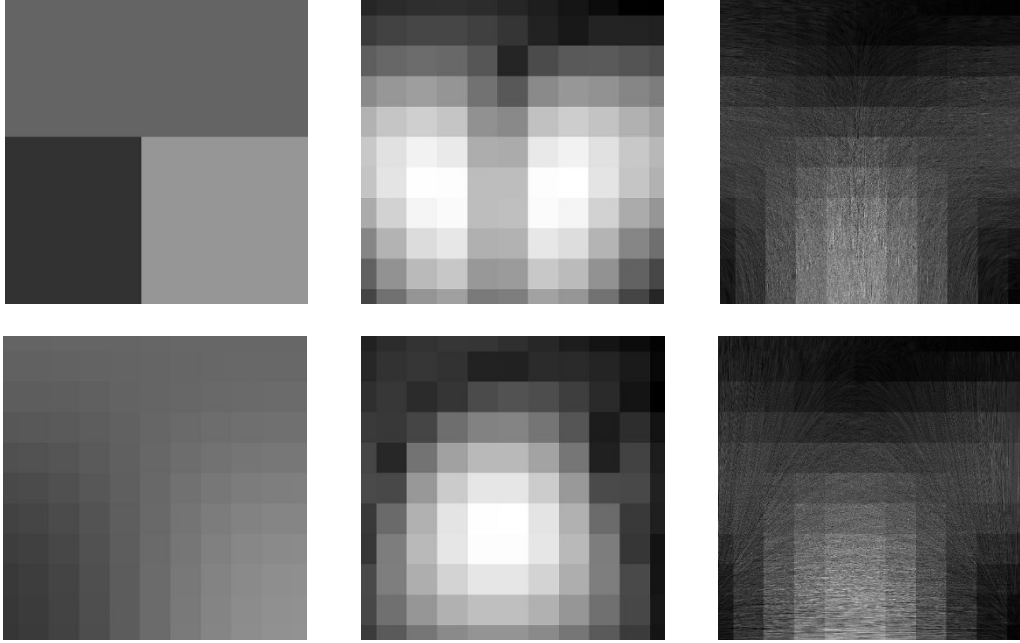

FIGURE 16: Curvature measures for a T-junction feature extracted from Figure 8. Upper left: original image, lower left: smoothed image, upper middle: isophote curvature measure $|m_i|$, and lower middle: flowline curvature measure $|m_f|$. The rightmost figures use the "spot noise" technique [46] to visualize the isophote vector field (upper) and the flowline (or gradient) vector field that is orthogonal to it (lower).

these curvature measures will follow after the discussion on surface curvatures.

The intersection curvature can also be used to define the *intersection curvature measure*

$$
\begin{aligned}
m(\phi) &= \kappa(\phi) \left(\frac{\sqrt{2}}{\sigma} l_1(\phi) \right)^3 \\
&= \frac{2\sqrt{2}}{\sigma^2} l_2(\phi_o) \cdot \frac{(l_1(\phi))^3}{\left[(\sigma/\sqrt{2})^2 + (l_1(\phi))^2 \right]^{3/2}},
\end{aligned}
\tag{97}
$$

which is approximately equal to $l_2(\phi) \cdot (2\sqrt{2}/\sigma^2)$ if the first-order directional derivative $l_1(\phi) \cdot \sqrt{2}/\sigma \gg 1$. The

intersection curvature measure is hence equal to the second-order directional derivative if the first-order directional derivative is sufficiently large.

7.4.2. Surface curvatures

Intersection curvatures are especially useful at positions where the gradient is zero, that is, for which $l_1(\phi) = 0$ for all directions ϕ, while $l_2(\phi) \neq 0$ for some angles ϕ, or equivalently $\sum_{m=0}^{2} l_{2-m,m}^2 \neq 0$. Such positions are called *singular points*, and are important landmarks in an image. They can be subdivided into minima, maxima, and saddle points.
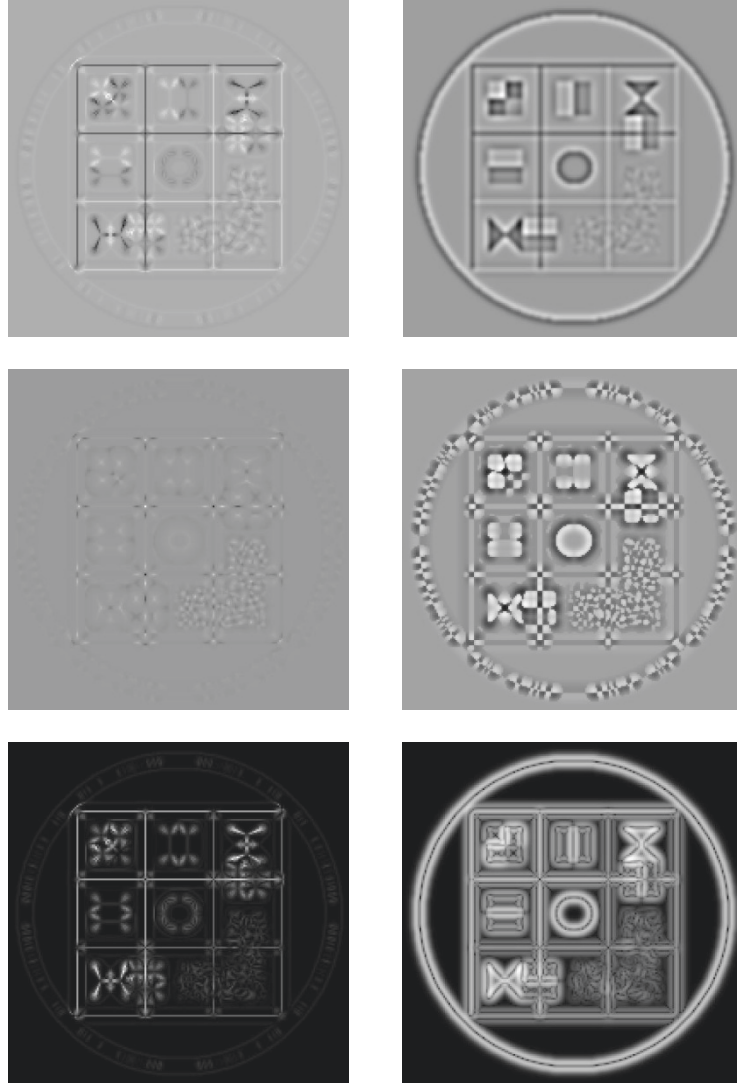
FIGURE 17: Curvature measures. Left from top to bottom: intersection curvatures κ_l, $\mathrm{sgn}(\kappa_g)|\kappa_g|^{1/2}$, and $|\kappa_r|^{1/2}$. Right from top to bottom: surface curvatures $\widetilde{\kappa}_l$, $\mathrm{sgn}(\widetilde{\kappa}_g)|\widetilde{\kappa}_g|^{1/2}$, and $|\widetilde{\kappa}_r|^{1/2}$.

In this case, the intersection curvatures are identical to the familiar *surface curvatures* [43], since translations in the image plane correspond one to one to translations in the tangent plane. For a singular point, we can arbitrarily select $\phi_d = 0$ as the reference angle, and obtain the following expression for the surface curvature in the direction ϕ:

$$
\begin{aligned}
\widetilde{\kappa}(\phi) &= \frac{2\sqrt{2}}{\sigma^2} l_2(\phi) \\
&= l_{20}^d \cdot \cos^2 \phi + 2l_{11}^d \cdot \cos\phi\sin\phi + l_{02}^d \cdot \sin^2 \phi \\
&= \frac{l_{20}^d + l_{02}^d}{2} + \sqrt{\left(l_{11}^d\right)^2 + \left(\frac{l_{20}^d - l_{02}^d}{2}\right)^2} \cdot \cos\left[2(\phi - \phi_c)\right],
\end{aligned}
\tag{98}
$$

where the direction ϕ_c of maximum surface curvature

$$
\widetilde{\kappa}_{\max} = \frac{l_{20}^d + l_{02}^d}{2} + \sqrt{\left(l_{11}^d\right)^2 + \left(\frac{l_{20}^d - l_{02}^d}{2}\right)^2}
\tag{99}
$$

satisfies

$$
\tan\left(2\phi_c\right) = \frac{2l_{11}^d}{l_{20}^d - l_{02}^d} = \frac{\sqrt{2}l_{11}}{l_{20} - l_{02}}.
\tag{100}
$$

The direction of minimum surface curvature

$$
\widetilde{\kappa}_{\min} = \frac{l_{20}^d + l_{02}^d}{2} - \sqrt{\left(l_{11}^d\right)^2 + \left(\frac{l_{20}^d - l_{02}^d}{2}\right)^2}
\tag{101}
$$

is $\phi_c + \pi/2$. The surface curvatures at a singular point hence belong to the range $[\widetilde{\kappa}_{\min}, \widetilde{\kappa}_{\max}]$, where the sum of the two

extreme curvatures is equal to the *Laplacian*, that is,

$$\widetilde{\kappa}_l = \widetilde{\kappa}_{\max} + \widetilde{\kappa}_{\min} = l_{20}^d + l_{02}^d = \text{trace } \mathcal{H}, \qquad (102)$$

and the product of the two extrema curvatures is equal to the *Gaussian curvature*

$$\widetilde{\kappa}_g = \widetilde{\kappa}_{\max} \cdot \widetilde{\kappa}_{\min} = l_{20}^d \cdot l_{02}^d - l_{11}^d \cdot l_{11}^d = \det \mathcal{H}. \qquad (103)$$

Both can be expressed in terms of the *Hessian matrix*

$$\mathcal{H} = \begin{bmatrix} l_{20}^d & l_{11}^d \\ l_{11}^d & l_{02}^d \end{bmatrix} = \frac{2}{\sigma^2} \begin{bmatrix} \sqrt{2}l_{20} & l_{11} \\ l_{11} & \sqrt{2}l_{02} \end{bmatrix}. \qquad (104)$$

A large absolute value for the Gaussian curvature indicates isolated singular points, since both extreme curvatures must have a large amplitude in this case. If the Gaussian curvature is positive, then both extreme curvatures have the same sign, which can only occur in case of an extremum point. The sign of the Laplacian $\widetilde{\kappa}_l$ allows to distinguish between minima ($\widetilde{\kappa}_l > 0$) and maxima ($\widetilde{\kappa}_l < 0$) in this case. A negative value for the Gaussian curvature indicates a saddle point.

If one of the extreme curvatures $\widetilde{\kappa}_{\max}$ or $\widetilde{\kappa}_{\min}$ is zero, then the Gaussian curvature $\widetilde{\kappa}_g$ is also zero. This occurs in case the singular points are not isolated but align to form extended ridges (line-like structures). The *ridge strength*

$$\widetilde{\kappa}_r^2 = \left| \widetilde{\kappa}_{\max}^2 - \widetilde{\kappa}_{\min}^2 \right|^2 = (l_{20}^d + l_{02}^d)^2 [4(l_{11}^d)^2 + (l_{20}^d - l_{02}^d)^2] \qquad (105)$$

can be used to find singular points for which there is a substantial difference between the amplitudes of both extreme curvatures. A large value for the ridge strength $\widetilde{\kappa}_r$, in combination with a small value for the Gaussian curvature $\widetilde{\kappa}_g$, is then evidence for the presence of ridge points.

All three mentioned surface curvature measures are illustrated in the right part of Figure 17.

Although the surface curvature measures only have a geometrical interpretation at singular points, they give nonzero responses for most nonuniform regions in an image. The gradient is hence needed alongside these measures in order to guide the interpretation. It is straightforward to derive the relationship

$$\kappa(\phi) = \frac{\widetilde{\kappa}(\phi)}{\left[1 + (l_{10}^d \cdot \cos(\phi - \phi_d))^2\right]^{3/2}} \qquad (106)$$

between intersection and surface curvature at positions where the gradient amplitude is l_{10}^d and the gradient direction is ϕ_d. Because the gradient amplitude is usually nonzero, except for uniform regions and singular points, the surface curvature differs from the intersection curvature in the sense that contributions for nonsingular points are being suppressed in the latter. This can be observed in Figure 17 by comparing the corresponding intersection and surface curvature measures on the left and right, respectively. This suppression of nonsingular points is expected to be beneficial for subsequent pattern matching or classification.

Note that, while edges are nonsingular points of the image itself, they are singular points of the image gradient, so that the focus on singular points can also be used to perform edge detection. Some intersection and surface curvature measures obtained from a gradient image are illustrated in Figure 18. A large value for the ridge strength κ_r, in combination with a small value for the Gaussian curvature κ_g, is then an evidence for the presence of edge points.

8. ANISOTROPIC NOISE REDUCTION

We finish by applying the partial image decomposition technique of Figure 5 to the case of the Hermite transformation in order to create an anisotropic noise reduction and contrast enhancement algorithm. More specifically, we will use a Hermite transformation up to order N, so that the index set $\mathcal{N}_o$ contains the Hermite coefficients $l_{n-m,m}$, for $n = 0, 1, \ldots, N$ and $m = 0, \ldots, n$. The mapping of these Hermite coefficients can be influenced by the following control variables.

(1) As in the case of residue-image processing, the average value l_{00} and the local energy E are used to determine the noise in the image. It was discussed earlier how the noise characteristics can be estimated from the histogram of the residue energy $E_r = E - l_{00}^2$ or residue amplitude $A_r = \sqrt{E_r}$.

(2) The Hermite coefficients up to order N can be mapped into a 1D and 2D component. Most often, the 1D energy $E_{2,1d}(\phi_o)$ in the first- and second-order Hermite coefficients is maximized to determine the direction ϕ_o of the local coordinate axes. The resulting adaptive Hermite transformation with coefficients $l_{n-m,m}(\phi_o)$, for $n = 0, 1, \ldots, N$ and $m = 0, \ldots, n$, allows to separate the energy up to order N into a 1D and 2D component, that is, $E_N = E_{N,1d}(\phi_o) + E_{N,2d}(\phi_o)$.

The mappings[2]

$$\widetilde{l}_{00}(\phi_o) = (1 - \kappa) \cdot l_{00}(\phi_o),$$

$$\widetilde{l}_{n0}(\phi_o) = (\kappa_1 - \kappa) \cdot l_{n0}(\phi_o),$$

$$\widetilde{l}_{n-m,m}(\phi_o) = (\kappa_2 - \kappa) \cdot l_{n-m,m}(\phi_o), \quad \text{for } m = 1, \ldots, n \qquad (107)$$

for the adaptive Hermite coefficients of orders $n = 0, 1, \ldots, N$ specify the adaptive algorithm, as depicted in Figure 5. The factors κ_1, κ_2, and κ are the amplification factors of the 1D Hermite component, the 2D Hermite component, and the residue component, respectively. The case of $\kappa_1 = \kappa_2 = \kappa$ corresponds to the residue-image processing algorithm of Section 4.

We of course need to make a specific choice for the amplification factors in order to realize our anisotropic noise reduction and/or contrast enhancement algorithm. Again, as in the case of residue-image processing, there are some heuristics involved in making this choice. A simple modification to the residue-image processing algorithm of Section 4 leads to

---

[2] Based on the argumentation in [14, 33], a mapping with κ_1 and κ_2 replaced by κ_1^n and κ_2^n may possibly constitute a better choice.
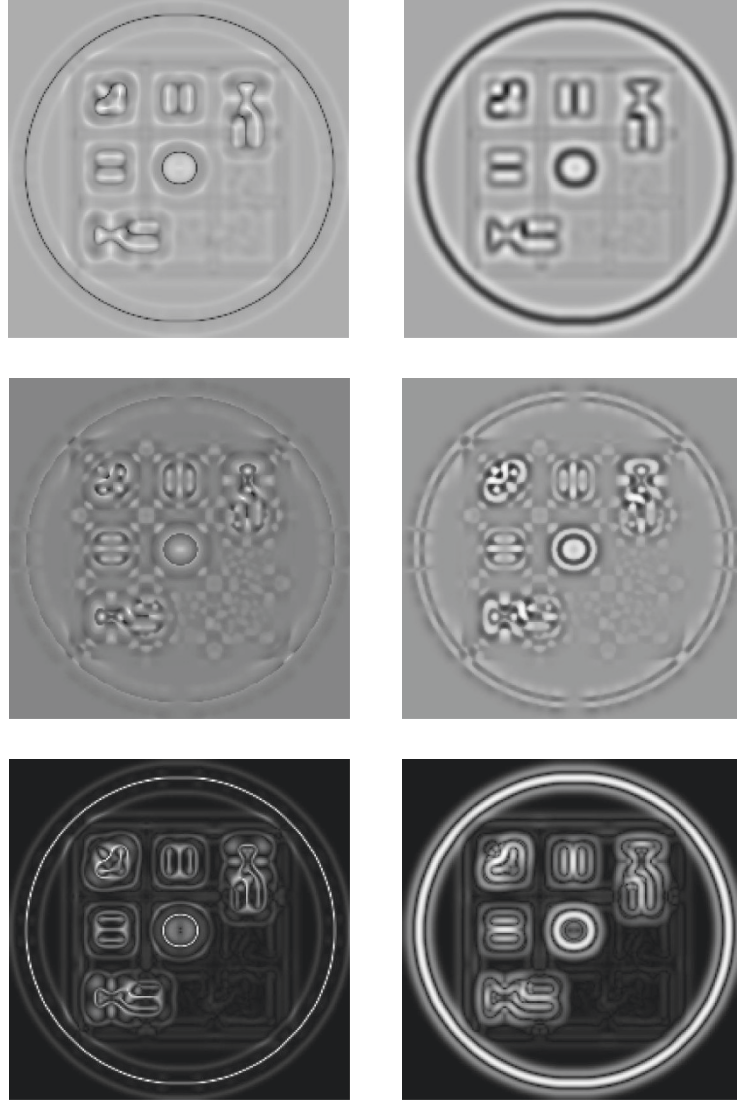
Figure 18: Curvature measures on gradient image. Left from top to bottom: intersection curvatures κ_l, $\mathrm{sgn}(\kappa_g)|\kappa_g|^{1/2}$, and $|\kappa_r|^{1/2}$. Right from top to bottom: surface curvatures $\tilde{\kappa}_l$, $\mathrm{sgn}(\tilde{\kappa}_g)|\tilde{\kappa}_g|^{1/2}$, and $|\tilde{\kappa}_r|^{1/2}$.

the following suggestion:

$$\kappa = \kappa_2 = s + (1-s)\frac{A_v}{A_r}, \qquad \kappa_1 = s_1 + (1-s_1)\frac{A_v}{A_r} \quad (108)$$

with

$$s_1 = s \cdot \left[1 - (1-\alpha) \cdot \left(\frac{E_{N,1d}(\phi_o)}{E_N}\right)^{\beta}\right], \qquad (109)$$

for $A_t \leq A_r \leq A_v$. The slope s_1 varies from s, in case $E_{N,1d}(\phi_o) = 0$, to $\alpha \cdot s$, with $0 \leq \alpha \leq 1$, for purely 1D features with $E_{N,1d}(\phi_o) = E_N$. The power $\beta \geq 0$ controls the steepness of this transition. No anisotropic processing is applied outside the range $[A_t, A_v]$, that is, $\kappa = \kappa_1 = \kappa_2 = 0$ for $A_r < A_t$ and $\kappa = \kappa_1 = \kappa_2 = 1$ for $A_r > A_v$.

In the upper-left corner of Figure 19, we show a detail of an original image. The results of anisotropic processing are illustrated in the bottom row of Figure 19 for some fairly arbitrary choices of the parameters. Specifically, $\alpha = 0$, $\beta = 0.2$ for the image in the lower left, and $\alpha = 0$, $\beta = 5$ for the image in the lower right. The value $A_t = 3$ is at twice the mode of the residue amplitude histogram, while the parameters controlling the amplification factors are fixed at $A_v = 30$ and $s = 0.5$. The image in the upper right-hand corner has been included for comparison. It is obtained using the residue-image processing algorithm of Section 4 (with the same parameter values) and corresponds to isotropic image processing. As can also be observed in the figure, the differences between isotropic and anisotropic filtering are often very subtle, so that a formal image-quality comparison [24]
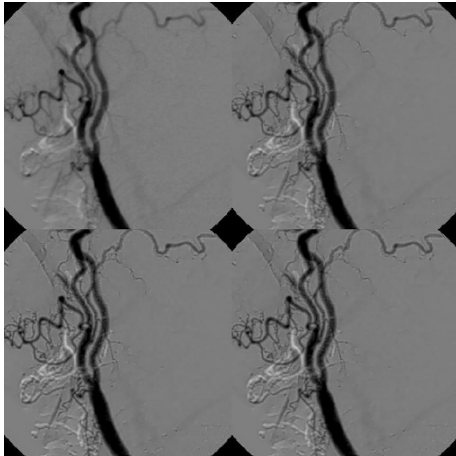
FIGURE 19: Results of anisotropic image processing with the Hermite transformation of order $N = 2$ and $\sigma = 1.5$ times the sampling distance. The original image is in the upper left-hand corner. The image in the upper right-hand corner is obtained through isotropic residue-image processing. The two images in the bottom row are obtained with the anisotropic image processing algorithm.

is probably needed to determine if the extra effort required for anisotropic filtering is truly worthwhile.

9. CONCLUSION

With this survey we have attempted to give an overview of the main characteristics of the Hermite transformation and their implications for applications. In summary, we can state that the characteristics of the two-dimensional Hermite transformation at a single scale (i.e., for one value of the Gaussian standard deviation σ) are by now well understood. In our view, the main challenges for the future lie in the extension to higher dimensions [26] and in a better exploitation of the characteristics of the Hermite transformation across multiple scales (i.e., more advanced than the simple pyramid extension used most of the time today). We need a better understanding of how the Hermite coefficients relate across scales, and how this relationship can be exploited more profitably for image coding, image processing, and image analysis [33, 42]. An improved link between the two alternative ways of interpreting multiscale Gaussian derivatives, that is, as analysis functions of the Hermite transformation or within the mathematical theory of scale space, may potentially also profit both approaches.

REFERENCES

[1] J.-B. Martens, "The Hermite transform—applications," *IEEE Transactions Acoustics, Speech, Signal Processing*, vol. 38, no. 9, pp. 1607–1618, 1990.

[2] J.-B. Martens, "The Hermite transform—theory," *IEEE Transactions Acoustics, Speech, Signal Processing*, vol. 38, no. 9, pp. 1595–1606, 1990.

[3] R. Young, "The Gaussian derivative theory of spatial vision: analysis of cortical cell receptive field line-weighting profiles," Tech. Report GMR-4920, General Motors Research Laboratories, Detroit, Mich, USA, 1985.

[4] J. L. Silván-Cárdenas and B. Escalante-Ramírez, "Image coding with a directional-oriented Hermite transform on a hexagonal lattice," in *Applications of Digital Image Processing XXIV*, A. G. Tescher, Ed., vol. 4472 of *Proceedings of SPIE*, pp. 528–536, San Diego, Calif, USA, July–August 2001.

[5] A. M. van Dijk and J.-B. Martens, "Feature-based image compression with steered Hermite transforms," in *Proceedings of IEEE International Conference on Image Processing (ICIP '96)*, vol. 1, pp. 205–208, IEEE, Lausanne, Switzerland, September 1996.

[6] A. M. van Dijk and J.-B. Martens, "Image representation and compression with steered Hermite transforms," *Signal Processing*, vol. 56, no. 1, pp. 1–16, 1997.

[7] B. Escalante-Ramírez and A. López-Caloca, "Image fusion with the Hermite transform," in *Proceedings of IEEE International Conference on Image Processing (ICIP '03)*, pp. 145–148, Barcelona, Spain, September 2003.

[8] B. Escalante-Ramírez, A. López-Caloca, and C. Zambrano-Gallardo, "Multiresolution fusion of remotely sensed images with the Hermite transform," in *Image and Signal Processing for Remote Sensing IX*, L. Bruzzone, Ed., vol. 5238 of *Proceedings of SPIE*, pp. 575–584, Barcelona, Spain, September 2004.

[9] B. Escalante-Ramírez, "Optic flow estimation using the Hermite transform," in *Applications of Digital Image Processing XXVII*, A. G. Tescher, Ed., vol. 5558 of *Proceedings of SPIE*, pp. 632–643, Denver, Colo, USA, August 2004.

[10] J. L. Silván-Cárdenas and B. Escalante-Ramírez, "Optic-flow information extraction with directional Gaussian-derivatives," in *Proceedings of IEEE 15th International Conference on Pattern Recognition (ICPR '00)*, vol. 3, pp. 3194–3197, Barcelona, Spain, September 2000.

[11] B. Escalante-Ramírez and J.-B. Martens, "Noise reduction in computerized tomography images by means of polynomial transforms," *Journal of Visual Communication and Image Representation*, vol. 3, pp. 272–285, 1992.

[12] B. Escalante-Ramírez, J.-B. Martens, and H. de Ridder, "Perceptually-based digital processing techniques for noise reduction in computed tomography images," in *Human Vision, Visual Processing, and Digital Display III*, B. E. Rogowitz, Ed., vol. 1666 of *Proceedings of SPIE*, pp. 288–299, San Jose, Calif, USA, February 1992.

[13] J.-B. Martens, "Deblurring digital images by means of polynomial transforms," *Computer Vision, Graphics and Image Processing*, vol. 50, no. 2, pp. 157–176, 1990.

[14] J.-B. Martens, "Adaptive image processing by means of polynomial transforms," in *Human Vision, Visual Processing, and Digital Display III*, B. E. Rogowitz, Ed., vol. 1666 of *Proceedings of SPIE*, pp. 276–287, San Jose, Calif, USA, February 1992.

[15] J.-B. Martens, "Adaptive contrast enhancement through residue-image processing," *Signal Processing*, vol. 44, no. 1, pp. 1–18, 1995.

[16] V. Kayargadde and J.-B. Martens, "Estimation of edge parameters and image blur using polynomial transforms," *CVGIP: Graphical Models and Image Processing*, vol. 56, no. 6, pp. 442–461, 1994.

[17] V. Kayargadde and J.-B. Martens, "An objective measure for perceived noise," *Signal Processing*, vol. 49, no. 3, pp. 187–206, 1996.

[18] V. Kayargadde and J.-B. Martens, "Estimation of perceived image blur using edge features," *International Journal of Imaging Systems and Technology*, vol. 7, no. 2, pp. 102–109, 1996.

[19] V. Kayargadde and J.-B. Martens, "Perceptual characterization of images degraded by blur and noise: model," *Journal of the Optical Society of America A*, vol. 13, no. 6, pp. 1178–1188, 1996.

[20] J.-B. Martens, "Local orientation analysis in images by means of the Hermite transformation," *IEEE Transactions Image Processing*, vol. 6, no. 8, pp. 1103–1116, 1997.

[21] C. J. Rivero-Moreno and S. Bres, "Spatio-temporal primitive extraction using Hermite and Laguerre filters for early vision video indexing," in *Proceedings of International Conference on Image Analysis and Recognition (ICIAR '04)*, pp. 825–832, Porto, Portugal, September–October 2004.

[22] C. J. Rivero-Moreno and S. Bres, "Texture feature extraction and indexing by Hermite filters," in *Proceedings of IEEE 17th International Conference on Pattern Recognition (ICPR '04)*, pp. 684–687, Cambridge, UK, August 2004.

[23] M. A. Georgeson and T. C. A. Freeman, "Perceived location of bars and edges in one-dimensional images: computational models and human vision," *Vision Research*, vol. 37, no. 1, pp. 127–142, 1997.

[24] J.-B. Martens, *Image Technology Design—A Perceptual Approach*, Kluwer Academic, Boston, Mass, USA, 2003.

[25] C. J. Rivero-Moreno and S. Bres, "Conditions of similarity between Hermite and Gabor filters as models of the human visual system," in *Proceedings of 10th International Conference on Computer Analysis of Images and Patterns (CAIP '03)*, pp. 762–769, Groningen, the Netherlands, August 2003.

[26] F. Brackx, N. de Schepper, and F. Sommen, "The higher dimensional Hermite transform—a new approach," *Complex Variables*, vol. 48, no. 3, pp. 189–210, 2003.

[27] S. G. Mallat, "Multifrequency channel decompositions of images and wavelet models," *IEEE Transactions Acoustics, Speech, Signal Processing*, vol. 37, no. 12, pp. 2091–2110, 1989.

[28] M. Vetterli and D. Le Gall, "Perfect reconstruction FIR filter banks: some properties and factorizations," *IEEE Transactions Acoustics, Speech, Signal Processing*, vol. 37, no. 7, pp. 1057–1071, 1989.

[29] I. Daubechies, *Ten Lectures on Wavelets*, SIAM, Philadelphia, Pa, USA, 1992.

[30] A. Rosenfeld, *Multiresolution Image Processing and Analysis*, Springer, Berlin, Germany, 1984.

[31] J. J. Koenderink and A. J. van Doorn, "Generic neighbourhood operators," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 14, no. 6, pp. 597–605, 1992.

[32] T. Lindeberg, *Scale-Space Theory in Computer Vision*, Kluwer Academic, Boston, Mass, USA, 1994.

[33] S. Makram-Ebeid and B. Mory, "Scale-space image analysis based on Hermite polynomials theory," in *Proceedings of 4th International Conference on Scale Space Methods in Computer Vision (Scale-Space '03)*, L. D. Griffin and M. Lillholm, Eds., vol. 2695 of *Lecture Notes in Computer Science*, pp. 57–71, Springer, Isle of Skye, UK, June 2003.

[34] J.-B. Martens, "Application of scale space to image coding," *IEEE Transactions on Communications*, vol. 38, no. 9, pp. 1585–1591, 1990.

[35] L. Florack, *Computational Imaging and Vision*, Kluwer Academic, Dordrecht, the Netherlands, 1997.

[36] E. Dubois, "The Sampling and reconstruction of time-varying imagery with application in video systems," *Proceedings of the IEEE*, vol. 73, no. 4, pp. 502–522, 1985.

[37] J. J. Koenderink, "The structure of images," *Biological Cybernetics*, vol. 50, no. 5, pp. 363–370, 1984.

[38] B. M. ter Haar Romeny, *Front-End Vision and Multi-Scale Image Analysis*, Kluwer Academic, Boston, Mass, USA, 2003.

[39] B. M. ter Haar Romeny, Ed., *Geometry-Driven Diffusion in Computer Vision*, Kluwer Academic, Dordrecht, the Netherlands, 1994.

[40] J. Weickert, B. M. ter Haar Romeny, and M. A. Viergever, "Efficient and reliable schemes for nonlinear diffusion filtering," *IEEE Transactions Image Processing*, vol. 7, no. 3, pp. 398–410, 1998.

[41] T. Lindeberg, "Edge detection and ridge detection with automatic scale selection," *International Journal of Computer Vision*, vol. 30, no. 2, pp. 117–156, 1998.

[42] T. Lindeberg, "Feature detection with automatic scale selection," *International Journal of Computer Vision*, vol. 30, no. 2, pp. 79–116, 1998.

[43] J. A. Thorpe, *Elementary Topics in Differential Geometry*, Springer, New York, NY, USA, 1979.

[44] W. T. Freeman and E. H. Adelson, "The design and use of steerable filters," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 13, no. 9, pp. 891–906, 1991.

[45] L. Meesters and J.-B. Martens, "A single-ended blockiness measure for JPEG-coded images," *Signal Processing*, vol. 82, no. 3, pp. 369–387, 2002.

[46] J. J. van Wijk, "Spot noise texture synthesis for data visualization," *Computer Graphics*, vol. 25, no. 4, pp. 309–318, 1991.

Jean-Bernard Martens studied at the State University of Gent, Belgium, where he received the Electrical Engineering degree in 1979. In 1983 he obtained his Ph.D. degree on a thesis entitled *Algorithms for the calculation of discrete convolutions and Fourier transformations*. He received an IEEE Paper Award for his paper entitled "Discrete Fourier transform algorithms for real valued sequences" in 1984. In that year, he also joined the Institute of Perception Research at Eindhoven University of Technology, where he became an Associate Professor in the field of visual psychophysics in 1991. He was an advisor for a number of Ph.D. projects on image quality, psychometrics, and image processing and coding. He (co-) authored several papers on these subjects and received the 1997 EURASIP Best Paper Award for a paper entitled "Image representation and compression with steered Hermite transforms" (authors: A. M. van Dijk and J. B. Martens). He also published the book *Image Technology Design—A Perceptual Approach* (Kluwer Academic Publishers) in 2003. He joined the Department of Industrial Design in 2003, where he is a full-time Professor in the field of visual interaction. His current research focuses on developing and testing new augmented-reality interaction styles in which working with images is an important component.