JOURNAL OF VISUAL COMMUNICATION AND IMAGE REPRESENTATION
Vol. 6, No. 2, June, pp. 154–163, 1995

Symmetry Interpretation of Complex Moments and the Local Power Spectrum*

JOSEF BIGÜN AND J. M. HANS DU BUF†

Signal Processing Laboratory, Swiss Federal Institute of Technology, EPFL-Ecublens, CH-1015 Lausanne, Switzerland

Received February 3, 1993; accepted November 30, 1993

Complex moments yield estimates of the local N-folded image symmetry because they give optimal solutions to the generalized line-fitting problem. Instead of computing them in the spatial domain, they can be computed in the local frequency domain because the local (Gabor) power spectrum is translation invariant (or varying slowly in practice) inside homogeneously textured regions. In other words, they make it possible to detect and discriminate linear, rectangular, hexagonal/triangular, etc., structures. When computed for different frequency bands they make it possible to distinguish between textures having different structures at different scales. Experimental results on texture segmentation confirm this. Furthermore, complex moments of the local power spectrum are shown to be related to Lie operators as well as prolate spheroidal functions, which allow for implementations that bypass the Gabor filtering. © 1995 Academic Press, Inc.

1. INTRODUCTION

Exploiting coordinate transformation theory, Bigün [2, 3] has shown that the detection of many higher order image primitives can be modeled as an abstract line detection problem. The simplest case is the detection of linear structures, which results in a local orientation estimate together with an associated certainty. Such an analysis can be performed at different scales, and the results indicate whether a texture is fine or coarse, isotropic or anisotropic [4, 5]. Other higher order primitives are spirals, circles, and T-junctions and can be dealt with by various differential operators [15, 16, 28]. All these approaches, which emphasize different types of invariances, such as shape and size, are related to Hoffman's Lie algebra of vision [22], and some evidence for the existence of polar and hyperbolic operators in the primate visual system has been found as well [19].

Complex moments of a function $\rho(x, y)$, which are defined by

$$I_{mn} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x + iy)^m (x - iy)^n \rho(x, y)\, dx\, dy \quad (1)$$

$$m, n = 0, 1, 2, \cdots,$$

are strongly connected to rotation invariant functions. Zernike [40] introduced them in relation with the design of nonlinear optical devices. In solving a rotation invariant differential equation, he obtained a set of radial and orthogonal polynomials also known as Zernike moments. Hu [25] derived real geometric moment invariants by using the algebraic theory of quantics as developed by Boole and Sylvester [14]. These invariants have often been applied to object recognition in the spatial domain, when the contour or skeleton of an object is well defined [13, 25]. Reddi [31] simplified the analytic derivation of geometric moment invariants by showing that they are related to complex moments, thus deriving what he called radial and angular moments. A related approach was proposed by Abu-Mostafa and Psaltis [1]. Their interpretation is based on the Fourier expansions of the spatial image along concentric circles. Other studies concerned with moment invariants (geometrical and/or Zernike invariants) in the spatial domain, for optical character recognition in particular, can be found in [27, 34, 37, 39].

Freeman and Saleh [17] took a different approach by combining complex moment magnitudes (Fourier–Mellin descriptors) in the spatial and in the frequency domain. Their approach relates to the ideas presented here, because it stresses the importance of features based on both domains. In addition, we will work toward a translation invariant and local image representation as opposed to the origin-dependent and global representation as employed by Abu-Mostafa and Psaltis [1]. We achieve this by computing complex moments on the basis of the local (Gabor) power spectrum for each position in an image. This local power spectrum is obtained by filtering an im-

* This work was supported by Thomson-CSF (France).

† Present address: Signals and Systems Group, UCEH—University of Algarve, Campus de Gambelas, 8000 Faro, Portugal.

1047-3203/95 $6.00

age with a complete set of complex Gabor filters, each with real and imaginary point spread functions, and computing the squared magnitude of all complex responses (the filtering is done in the frequency domain of course). Therefore, the local power spectrum is translation invariant inside homogeneously textured regions, provided that the Gabor filters are tuned to frequencies not too high. In other words, they may not act like line and edge detectors [11].

Below we will show that complex moments computed in the local frequency domain can yield a local and translation invariant description of an image. This new and elegant interpretation completes in the mathematical sense the linear symmetry description of a neighborhood [4] and introduces N-folded symmetries. Furthermore, we will link complex moments to the local or Gabor power spectrum, which is very "en vogue" in texture analysis, e.g. [26, 30, 32], but with the goal of deriving semantically meaningful image features. The generalization of linear symmetry to N-folded symmetries, for which we present only one application here, namely texture segmentation, also completes other theories like steerable filters used for example in orientation estimation [18, 20, 29]. The steerable-filter approach for orientation detection is limited to linear symmetry cases, whereas complex moments are capable of addressing general symmetries, thereby obtaining not only the symmetry information but also the orientation information. This makes it possible to construct operators which can detect very specific structures like a rectangular structure having a given orientation.

We emphasize that our main concern here is with theory. We will focus on the mathematical proof of the generalized line-fitting problem by means of complex moments, the interpretation of the complex moments, and the theoretical relations with Lie operators as well as prolate spheroidal functions. The latter methods allow for different implementations which bypass the Gabor filtering stage.

Furthermore, we will only illustrate one application, texture segmentation, in order to underline the power of the complex moment approach. However, we will not go into details because a second study is concerned with extensive tests and comparisons with alternative methods [7]. These include different sets of real moments of the local power spectrum which lack a clear semantical interpretation. Although experimental results are quite comparable, the advantage of the complex moment interpretation makes this approach preferable. In addition, the same test images and the same processing scheme have been utilized in other studies in which other texture feature extraction methods have been compared [4, 6]. The results obtained were clearly inferior.

The rest of this paper is organized as follows. The solution of the line-fitting problem utilizing complex moments is presented in Section 2. This problem is generalized to N-folded symmetries in Section 3. Important aspects of the application of complex moments to the Gabor power spectrum are discussed in Section 4 and illustrated by an application, namely unsupervised texture segmentation. The relations between complex moments of the local power spectrum and Lie operators as well as prolate spheroidal functions are studied in Section 5.

2. THE OPTIMAL LINE-FITTING PROCESS

Suppose that we want to fit a line to a two-dimensional, continuous, real, and nonnegative function $\rho(\bar{x})$, $\bar{x} \in E_2$, in a least-squares error (LSE) sense. Furthermore, we are only interested in fitting a line which goes through the origin, see also Fig. 1 (right). Here E_2 represents the two-dimensional Euclidean space, and the error to be minimized over $\bar{k}$, with $\|\bar{k}\| = 1$, is

$$e(\bar{k}) = \int_{\bar{x} \in E_2} \|\bar{x} - (\bar{k}^t \bar{x})\bar{k}\|^2 \rho(\bar{x})\, d\bar{x}, \tag{2}$$

where $\|\bar{x} - (\bar{k}^t \bar{x})\bar{k}\|$ is the distance of a point $\bar{x}$ to the line through the unit vector $\bar{k}$. This problem can be solved by using straightforward differential calculus, but we will present a solution based on complex moments because this will be exploited in a more general sense later on. From (2) we obtain by algebraic manipulations

$$e(\bar{k}) = \int_{\bar{x} \in E_2} \|\bar{x}\|^2 \rho(\bar{x})\, d\bar{x} - e_1(\bar{k}). \tag{3}$$

Here $e_1(\bar{k})$ is defined as

$$e_1(\bar{k}) = \int_{\bar{x} \in E_2} |\bar{k}^t \bar{x}|^2 \rho(\bar{x})\, d\bar{x}, \tag{4}$$

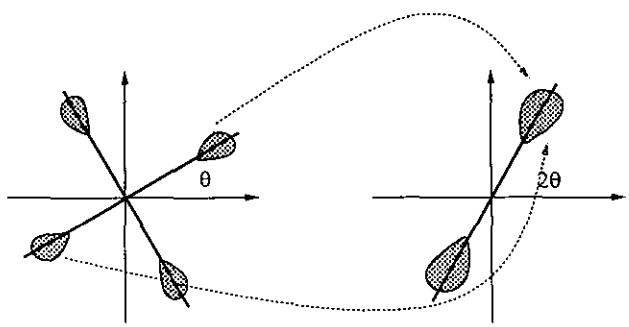

FIG. 1. The cross-fitting problem is solved by applying a coordinate transform first, followed by a line-fitting in the new coordinate system.

and the translated problem is to find

$$\max_{\|\bar{k}\|=1} e_1(\bar{k}). \tag{5}$$

Before solving this problem we introduce the following notation. We had $\bar{x} \in E_2$, that is, $\bar{x}$ is a vector with two real components:

$$\bar{x} = (x_1, x_2)^t. \tag{6}$$

Then $\hat{x}$ is the complex number which corresponds to the point represented by $\bar{x}$:

$$\hat{x} = x_1 + ix_2. \tag{7}$$

Conversely, let $\hat{x}$ be a complex number. Then $\bar{\hat{x}}$ is a vector in E_2 which corresponds to $\hat{x}$:

$$\bar{\hat{x}} = (\mathrm{Re}(\hat{x}), \mathrm{Im}(\hat{x}))^t. \tag{8}$$

As a result, the symbol at the top of a variable determines the final interpretation, e.g., $\bar{\hat{x}}$ and $\hat{\bar{x}}$ represent a vector and a complex scalar originating from a complex scalar and a vector, respectively. The basic motivation for this notation is that we want to use the exponentiation, e.g., squaring properties of the complex number field while exploiting the geometric projection properties of the vector fields, at the same time.

The scalar product of two vectors $\bar{x}$ and $\bar{y}$ can be computed in the complex number field by using

$$\bar{x}^t\bar{y} = (\hat{\bar{x}}^*\hat{\bar{y}} + \hat{\bar{x}}\hat{\bar{y}}^*)/2 = \mathrm{Re}(\hat{\bar{x}}^*\hat{\bar{y}}). \tag{9}$$

Using (9), $e_1(\bar{k})$ of (4) is rewritten as

$$\begin{aligned} e_1(\bar{k}) &= \int_{\bar{x}\in E_2} (\mathrm{Re}(\hat{\bar{k}}^*\hat{\bar{x}}))^2\rho(\bar{x})\,d\bar{x} \\ &= \frac{1}{4}\int_{\bar{x}\in E_2} [(\hat{\bar{k}}^*\hat{\bar{x}})^2 + (\hat{\bar{k}}\hat{\bar{x}}^*)^2 + 2(\hat{\bar{k}}\hat{\bar{k}}^*)\hat{\bar{x}}\hat{\bar{x}}^*]\rho(\bar{x})\,d\bar{x} \\ &= \frac{1}{4}\int_{\bar{x}\in E_2} [(\hat{\bar{k}}^2)^*(\hat{\bar{x}})^2 + ((\hat{\bar{k}}^2)^*(\hat{\bar{x}}^2)^* + 2|\hat{\bar{x}}|^2]\rho(\bar{x})\,d\bar{x} \\ &= \frac{1}{2}\int_{\bar{x}\in E_2} |\hat{\bar{x}}|^2\rho(\bar{x})\,d\bar{x} + \frac{1}{2}\,\mathrm{Re}\left(\hat{\bar{k}}^2\int_{\bar{x}\in E_2}\hat{\bar{x}}^2\rho(\bar{x})\,d\bar{x}\right), \end{aligned} \tag{10}$$

where $\hat{\bar{k}}\hat{\bar{k}}^* = |\hat{\bar{k}}|^2 = 1$ has been applied to obtain the third line. Hence, our maximization problem is once more translated into another one, namely to maximize $e_2(\bar{k})$ which is defined by

$$e_2(\bar{k}) = \mathrm{Re}(\hat{\bar{k}}^2\hat{z}) = \overline{(\hat{\bar{k}}^2)}^t\,\bar{\hat{z}}, \tag{11}$$

where $\hat{z}$ is a complex moment (1) which is independent of $\bar{k}$:

$$\hat{z} = I_{20} = \int_{\bar{x}\in E_2} \hat{\bar{x}}^2\rho(\bar{x})\,d\bar{x}. \tag{12}$$

Obviously, the projection of $\bar{\hat{z}}$ onto the same direction as itself is maximum among all other projections:

$$\hat{\bar{k}}^2_{\max} = \frac{\hat{z}}{|\hat{z}|}. \tag{13}$$

Consequently, we obtain the optimal error

$$e(\bar{k}_{\min}) = \frac{1}{2}\int_{\bar{x}\in E_2} |\hat{\bar{x}}|^2\rho(\bar{x})\,d\bar{x} - \frac{1}{2}|\hat{z}|. \tag{14}$$

In a similar way, a direction $\bar{k}_{\max}$ which maximizes $e(\bar{k})$ of (2), thereby minimizing e_2 of (11), can be found. This yields the error

$$e(\bar{k}_{\max}) = \frac{1}{2}\int_{\bar{x}\in E_2} |\hat{\bar{x}}|^2\rho(\bar{x})\,d\bar{x} + \frac{1}{2}|\hat{z}|. \tag{15}$$

By using (14) and (15) we also note that the sum of the maximum and minimum error is a complex moment, namely I_{11}. Moreover, we note that if $\bar{k}_{\min}$ minimizes (2), so does $-\bar{k}_{\min}$, and the solution is unique up to $\pm\pi$. The direction given by I_{20} automatically solves this ambiguity and we represent the optimal $\bar{k}$ by $\hat{\bar{k}}^2$. The line-fitting process described up to now constitutes the proof of the following theorem:

THEOREM 1. *The second-order complex moments I_{11} and I_{20} determine an optimal line fitted to the function $\rho(\bar{x})$ in the LSE sense. The magnitudes of these moments are related to the absolute minimum and maximum error as*

$$\begin{aligned} |I_{20}| &= e(\bar{k}_{\max}) - e(\bar{k}_{\min}) \\ I_{11} &= e(\bar{k}_{\max}) + e(\bar{k}_{\min}) \end{aligned}$$

and the minimum and maximum error lines are determined by

$$\hat{\bar{k}}^2_{\min} = \frac{I_{20}}{|I_{20}|} \quad \text{and} \quad \hat{\bar{k}}^2_{\max} = -\frac{I_{20}}{|I_{20}|}. \tag{16}$$

3. FITTING AN OPTIMAL CROSS AND BEYOND

We will now be interested in fitting a cross to $\rho(\bar{x})$ in the LSE sense. The error function associated with this problem is not as straightforward to formulate as in the

previous section. The approach taken here is first to map ρ to another function through a two to one transformation, and then to apply the line-fitting problem in the new domain. After presenting our solution to this problem, we will generalize these results by means of a theorem which includes both the line- and cross-fitting processes as special cases. The proof of the theorem will not explicitly be carried out since it is a straightforward extension of the method described here. However, to facilitate further investigations, we will be slightly redundant and write 4/2 instead of just 2. We will show that I_{mn} with $m - n = 4$ (fourfolded symmetry) fits a cross to the function $\rho(\bar{x})$ in the LSE sense. Without loss of generality we consider I_{40} throughout this section. We first perform a Cartesian to polar transformation of the integral defining I_{40}:

$$\begin{aligned} I_{40} &= \int_{\bar{x}\in E_2} \hat{\bar{x}}^4 \rho(\bar{x})\, d\bar{x} \\ &= \int_0^{2\pi}\int_0^{\infty} r^4 \exp(i4\varphi)\rho(r,\varphi) r dr d\varphi. \end{aligned} \tag{17}$$

Using the transformation

$$\varphi_1 = \frac{4}{2}\varphi, \tag{18}$$

the above integral can be transformed once more:

$$\begin{aligned} I_{40} &= \sum_{j=0}^{4/2-1} \int_0^{2\pi(4/2)} \int_0^{\infty} r^4 \exp(i2\varphi_1)\rho\left(r, \frac{\varphi_1}{4/2} + j\frac{2\pi}{4/2}\right) \\ &\quad \frac{1}{4/2}\, r dr d\varphi_1 \\ &= \int_0^{2\pi(4/2)} \int_0^{\infty} [r\exp(i\varphi_1)]^2 \left[\sum_{j=0}^{4/2-1} r^{4-2}\rho\left(r, \frac{\varphi_1}{4/2} + j\frac{2\pi}{4/2}\right)\right] \\ &\quad \frac{1}{4/2}\, r dr d\varphi_1. \end{aligned} \tag{19}$$

For a further simplification we need to show that the sum in the integrand is 2π periodic:

$$\begin{aligned} &\sum_{j=0}^{4/2-1} r^{4-2}\rho\left(r, \frac{\varphi_1 + 2\pi}{4/2} + j\frac{2\pi}{4/2}\right) = \\ &\sum_{j=0}^{4/2-1} r^{4-2}\rho\left(r, \frac{\varphi_1}{4/2} + (j+1)\frac{2\pi}{4/2}\right) = \\ &\sum_{j=0}^{4/2-1} r^{4-2}\rho\left(r, \frac{\varphi_1}{4/2} + j\frac{2\pi}{4/2}\right). \end{aligned} \tag{20}$$

The last equality is obtained by observing that the integer multiples of the angles $2\pi/4/2$ constitute a group, the cyclic group, with angle addition (circular) being the group addition operation. Returning to (19), and using the 2π periodicity when splitting the integration interval $[0, 2\pi\frac{4}{2}]$, the factor 1/4/2 is eliminated so that we obtain

$$I_{40}(\rho) = \int_{\bar{x}\in E_2} \hat{\bar{x}}^2 \rho_1(\bar{x})\, d\bar{x} = I_{20}(\rho_1), \tag{21}$$

where

$$\rho_1(\bar{x}) = \sum_{j=0}^{4/2-1} |\hat{\bar{x}}|^{4-2}\rho\left(|\hat{\bar{x}}|, \frac{\arg \hat{\bar{x}}}{4/2} + j\frac{2\pi}{4/2}\right). \tag{22}$$

Similarly, it can be shown that

$$I_{22}(\rho) = I_{11}(\rho_1). \tag{23}$$

Hence, according to Theorem 1, $\arg(I_{40})$ determines the orientation of an optimal line fitted to the function $\rho_1(\bar{x})$, and $|I_{40}|$ is the difference between the associated largest and smallest errors. In (22), we note that ρ_1 is obtained by adding the values of ρ at 4/2 points and by weighting the sum with a radial function. These points participating in the addition lie on lines through the origin with equidistant angle separations of $2\pi/(4/2)$.

Hence, $I_{40}(\rho)$ first implicitly maps all points possessing the same fourfolded symmetry (points on the same cross) to a line in an additive manner. Then it fits a line to this new function in the LSE sense; see Fig. 1. The error is zero if and only if the points where $\rho(\bar{x})$ are nonzero possess a fourfolded symmetry (points should lie on a single ideal cross). A straightforward extension of the cross-fitting process to N-folded symmetry fitting gives us the following theorem:

THEOREM 2. *A pair of complex moments $I_{mn}(\rho)$ and $I_{(m+n)/2,(m+n)/2}(\rho)$, with a given $m - n \neq 0$, determines an optimal fit of a set of lines possessing $(m - n)$-folded symmetry to a function ρ. A set of lines has $(m - n)$-folded symmetry if the lines pass through the origin and any of them can be obtained from another by an integer multiple of $2\pi/(m - n)$ rotations. The fitting process is represented by*

$$\min_{\|\bar{k}_1\|=1} e(\bar{k}) = \min_{\|\bar{k}_1\|=1} \int_{\bar{x}\in E_2} \|\bar{x} - (\bar{k}_1^t\bar{x})\bar{k}_1\|^2 \rho_1(\bar{x})\, d\bar{x},$$

where

$$\rho_1(\bar{x}) = |\hat{\bar{x}}|^{m+n-2} \sum_{j=0}^{(m-n)/2-1} \rho\left(|\hat{\bar{x}}|, \frac{\arg \hat{\bar{x}}}{(m-n)/2} + j\frac{2\pi}{(m-n)/2}\right),$$

and $\bar{k}_1$ is a vector implicitly representing the orientation of the symmetry, $\hat{\bar{k}}^{m-n}$, through the relation $\hat{\bar{k}}_1^2 = \hat{\bar{k}}^{m-n}$.

The minimum and maximum errors are given by

$$|I_{mn}| = e(\bar{k}_{\max}) - e(\bar{k}_{\min}) \tag{24}$$

$$I_{(m+n)/2,(m+n)/2} = e(\bar{k}_{\max}) + e(\bar{k}_{\min}), \tag{25}$$

whereas the minimum and maximum error orientations of the symmetric set of the lines are determined by

$$\hat{\bar{k}}^{m-n}_{\min} = \frac{I_{m-n}}{|I_{m-n}|} \quad \textit{and} \quad \hat{\bar{k}}^{m-n}_{\max} = -\hat{\bar{k}}^{m-n}_{\min}. \tag{26}$$

4. THE LOCAL POWER SPECTRUM AND ITS REPRESENTATION BASED ON SYMMETRIES

Theorem 2 can be applied to the spatial domain, but also to the local frequency domain. The great advantage of considering the local power spectrum is that it is translation invariant within homogeneous regions, i.e., local phase information which makes it possible to discriminate between textures consisting of differently shaped lines and edges is neglected, e.g., [11]. In other words, only the geometrical structures are taken into account, irrespective of the shape of the structure's components. As already discussed in the Introduction, translation invariance requires that the filters employed are tuned to frequencies not too high. Therefore, when there is a texture with one distinct orientation (linear or twofolded symmetry) around an inspected point, the power spectrum will be concentrated to a line. When there is a texture with two, mutually orthogonal, directions (rectangular or fourfolded symmetry) the power spectrum will be concentrated to a cross. It is similar for hexagonal/triangular (sixfolded symmetry) and octagonal structures (eightfolded symmetry). The arguments of the complex moments give the orientation of the estimated N-folded symmetry, whereas the magnitudes give measures of the estimation quality, that is, certainties.

At this point we note that if the complex moments I_{40} and I_{60} of a local power spectrum with an energy concentration having only twofolded symmetry is computed, the magnitude responses will be high. However, these magnitude responses will be lower than those in the cases in which the spectrum shows four- and six-folded symmetries. The reason for this is that 2 is a factor of 4 and 6; hence, such a concentration "leaks" to four- and sixfolded symmetries. The converse is not true, which allows complex moments to discriminate between these symmetries. In the ideal case, a sixfolded symmetric energy concentration gives a zero response for complex moments with $m - n = 4$ and 2. Likewise, a fourfolded symmetric concentration yields a zero response for $m - n = 2$ and 6. In practice, the discriminative power of complex moments will be less when the energy in the measured local power spectrum is spread. A simple consequence of the Cauchy-Schwartz inequality is Heisenberg's uncertainty principle, $W_x W_\omega \geq \frac{1}{2}$; see [21], where W_x and W_ω are the effective widths of a function as measured by the square root of the variance in the spatial and frequency domains. Hence, one can increase the orientation selectivity of the filters by taking very elongated point spread functions, but this leads of course to averaging across region boundaries.

Consider $\rho(\omega_{ri}, \omega_{\varphi j})$ when ω_{ri} is fixed to ω_{rk}. In a Gabor decomposition scheme with six orientation channels ($j = 0, \ldots, 5$) this subset of the local power spectrum corresponds to the responses of all six filters with different orientations tuned to the same radial frequency. Assume that only this subset is used in moment computations. There are four complex moments of order 6: I_{33}, I_{42}, I_{51}, and I_{60}, since $I_{mn} = I^*_{nm}$. There are no odd symmetries because the power spectrum is even symmetric. The sequence $\{\rho(\omega_{rk}, j(2\pi/12))\}_{j=0}^{11}$ is the power spectrum covering the entire spectrum over a closed circle. The inverse DFT of this sequence is given by

$$p_l = \sum_{j=0}^{11} \exp\left(ijl\,\frac{2\pi}{12}\right)\rho\left(\omega_{rk}, j\,\frac{2\pi}{12}\right). \tag{27}$$

By keeping in mind that the DFT causes $p_l = p^*_{12-l}$, it can be seen that $p_{m-n}\omega^6_{rk} = I_{mn}$ with $m + n = 6$. Since ω_{rk} is independent of ρ and is known a priori, the I_{mn}'s over a frequency band represent the inverse DFT. Hence, I_{33}, I_{42}, I_{51}, and I_{60} are actually the only moments required in order to uniquely represent the power spectrum for a given frequency band. This set of moments corresponds to six independent real variables, because only I_{42} and I_{51} are really complex. This can be seen by evaluating p_l for $l = 0$ and 6. When such a computation in terms of complex moments is carried out for each radial frequency ω_{ri}, one has actually applied an orthogonal transformation to the entire local power spectrum, since the Fourier transform is orthogonal and the moments cover nonoverlapping frequency bands. The generalization of the relation between the Fourier transform and the complex moments for Gabor decompositions with arbitrary numbers of filters is straightforward and therefore omitted here.

In order to visualize the effect of computing complex moments in the local power spectrum we have applied sixth-order moments (I_{33}, I_{42}, I_{51}, and I_{60}). Figures 2 and 3 show a test image with structured textures from the Brodatz album [9] together with the magnitudes of some moments in different frequency channels. The Gabor decomposition was similar to that exploited before in other contexts like the analysis of local phase information [10], namely a filter rosette in the frequency domain with five frequency bands in octaves and six equidistant orientations. In order to address the textures in the test image

we make use of the following coding:

T4	T7	T6	T1
T2	T1	T4	T5
T4	T3	T2	T7
T2	T5	T6	T3

The upper-right image in Fig. 2 is $|I_{42}|$ in frequency channel 1 (the lowest frequency band). Bright areas correspond to large magnitudes, and thus to high confidence linear symmetries. Texture T1, which roughly has this property, is highlighted, but the boundary uncertainty is large because of the low frequency of this channel. The bottom-left image is $|I_{42}|$ in frequency channel 2, which responds to the linear substructure of T2. The coarse linear pattern in the same texture excites $|I_{51}|$ in the same frequency channel, which is illustrated in the bottom-right image (recall that 2 is a factor of 4).

The upper-right image in Fig. 3 corresponds to $|I_{60}|$ in frequency channel 3. This symmetry type represents hexagonal as well as triangular structures. The textures T5 and T6 have this property. The response to T2 in this image can be explained by the fact that this texture includes substructures with two-folded symmetries (2 is a factor of 6). It is worth noting that the converse is not true, that is, the response of $|I_{42}|$ to textures with a hexagonal structure is low (see upper-right image of Fig. 2). The bottom-left image in Fig. 3 shows $|I_{51}|$ in frequency channel 4. Due to their rectangular structure, T4 and T3 respond in this image. The modulus of I_{42} in frequency channel 5 is shown in the bottom-right image. Here most of the texture boundaries are visible because they do not respond to this symmetry. The reason for this is that the mixture of two power spectra at boundaries normally lacks a single orientation and thus any linear symmetry (the uncertainty principle again). In conclusion, we have shown that the different moments in different frequency bands respond to specific structures, that some symmetry orders leak to multiple moments, but that the entire set of moments makes it possible to discriminate all textures. This can be exploited in texure segmentation; see below. However, if region-based segmentation methods are to be employed, we can expect some artifacts at region boundaries.

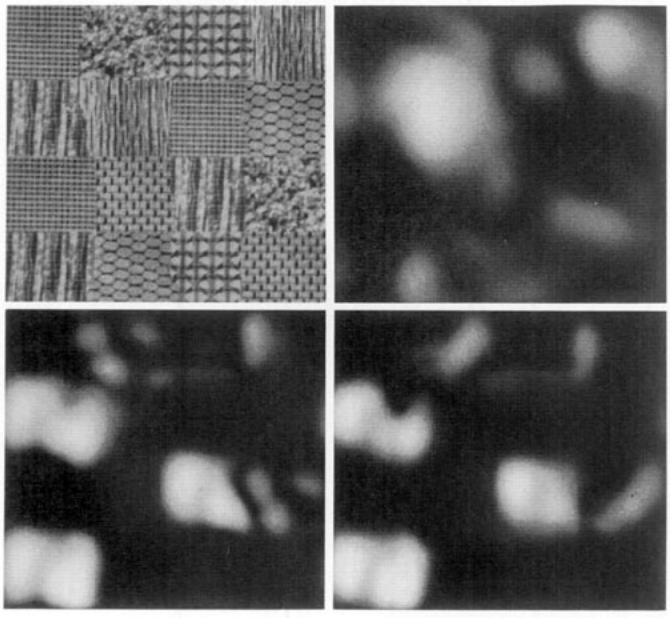

FIG. 2. Upper left, input image; upper right, $|I_{42}|$ in frequency channel 1 (lowest frequency); lower left, $|I_{42}|$ in frequency channel 2; lower-right, $|I_{51}|$ in frequency channel 2.

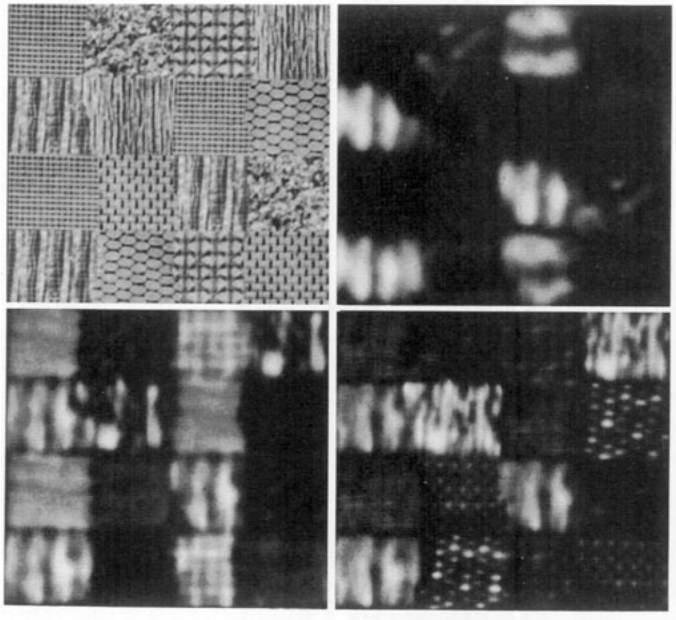

FIG. 3. Upper left, input image; upper right, $|I_{60}|$ in frequency channel 3; lower left, $|I_{51}|$ in frequency channel 4; lower right, $|I_{42}|$ in frequency channel 5.

One might think that the arguments of the complex moments which represent $\hat{k}^2$, $\hat{k}^4$, . . . according to Theorem 2 need to be divided by 2, 4, . . . for a straightforward representation of the orientations. Likewise, the moduli of these moments do not provide for the minimum error directly. Although it is simple to compute this by solving the minimum error in (37)–(38), we will argue that for pattern recognition purposes the complex number representation of I_{mn} has advantages which make it possible to circumvent the following three problems. First, the *continuity* of the orientations is preserved in the sense that two orientations (orientation of the lines, crosses, etc.) which differ a small amount also differ a small

amount in their numerical representations. That is, a number arbitrarily close to 0 is not arbitrarily close to 2π, whereas the corresponding physical angles are; see [20]. In the Cartesian representation the complex numbers are continuous with respect to changes in their arguments, including the angle of 0. Second, the factor $m - n$ makes the representation of the orientation of the symmetry unique. A line with an angle of $\pi/6$ is the same as the line with $(\pi/6) + \pi$, as are a cross with a (rotation!) angle of $\pi/6$ and the crosses with angles of $(\pi/6) + (\pi/2)$, $(\pi/6) + 2(\pi/2)$, and $(\pi/6) + 3(\pi/2)$. However, eliminating the factor $m - n$ would discard the physical equivalence of these cross rotation angles, causing what is called the *group representation* problem. Since the arguments of the complex moments actually correspond to $\hat{k}^{m-n}$, such an ambiguity does not exist because all angles are multiplied by $m - n$. Third, knowing the value of $e(\bar{k}_{\min})$ alone is not sufficient to judge the quality of the estimate; one must know whether this error is large or small (the *range* problem). The comparison with the worst case, i.e., $e(\bar{k}_{\max}) - e(\bar{k}_{\min})$, provides a means to assess the quality. Again, I_{mn} already represents the difference of the errors through its magnitude. Alternative quality measures can be found easily. As an example we mention the energy independent ratio

$$\frac{e(\bar{k}_{\max}) - e(\bar{k}_{\min})}{e(\bar{k}_{\max}) + e(\bar{k}_{\min})} = \left| \frac{I_{mn}}{I_{(m+n)/2,(m+n)/2}} \right|, \tag{28}$$

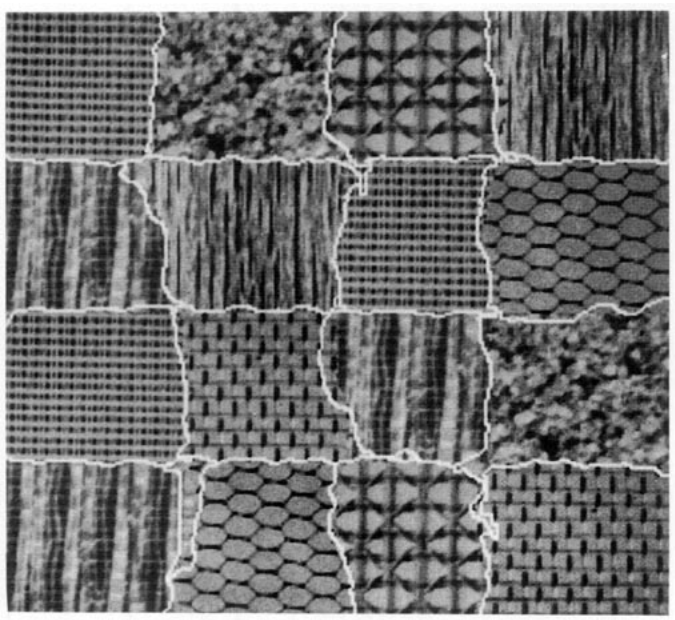

FIG. 4. Unsupervised quadtree segmentation by applying octave-band weighted moments to structured textures with two-, four-, and six-folded symmetries. From [7].

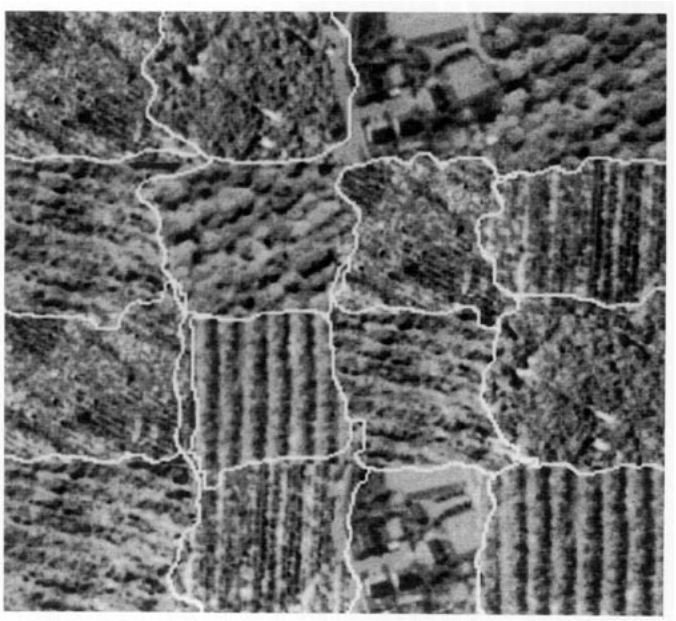

FIG. 5. Segmentation of aerial textures. From [7].

which is always in the interval [0, 1] and attains the endpoints 0 and 1 if the quality of the fit is totally uncertain and totally certain, respectively. The continuity, group representation, and range problems might easily imperil the performance of, e.g., a clustering method which could follow the extraction of orientation and certainty features, if an adequate representation of these features is lacking. By using complex moments in the Cartesian representation one can avoid these three problems.

Finally, Figs. 4 and 5 show the result of one application, namely unsupervised image segmentation. These results were obtained by complex moments with octave-band frequency weighting. This weighting is employed in order to compute the moments in different frequency bands, rather than summing over the entire local power spectrum. Such a procedure also yields a scale selectivity. Then a local Karhunen–Loève transform (LKL) (see [6]) was applied to the 14 real and imaginary parts of the complex moment images, and the reduced feature set was segmented by means of a quadtree method. For full details, and a complete experimental test which includes real moments as well, see [7]).

Here we note that the applied LKL transform is orthogonal; i.e., it preserves distances in the feature space. The complex moments of the Gabor power spectrum constitute a linear transform which is in general not orthogonal. Consequently, the complex-moment transform cannot be bypassed by applying the LKL transform directly to the Gabor power spectrum, since not all linear transforms are reachable by orthogonal transforms. However, it should be mentioned that orthogonal complex moments exist, namely the Zernike moments [40].

From Figs 4 and 5 we see that complex moments provide for very strong texture features, yielding good segmentation results. Not only for highly structured textures with N-folded symmetries but also for aerial textures which are less structured.

5. OTHER INTERPRETATIONS AND IMPLEMENTATIONS

In this section we investigate the precise physical meaning of the line-fitting, cross-fitting, etc. processes explored above. First, we relate the error function to some differential operators which offer not only another theoretical derivation of the same method but also another implementation in which the Gabor decomposition can be circumvented. Then we will do the same by considering prolate spheroidal functions.

5.1. *Relation to Lie Groups of Transformations*

By using the theory presented, a conceptive relation between complex moments of the local spectrum and some of the Lie transformation groups can be established. This relation is particularly strong because it gives an interpretation of the error function, which was originally formulated in the Gabor space, directly in the spatial domain by means of physical operations, that is, small translations. In order to analyze the significance of complex moments computed from the local power spectrum, but now in the spatial domain, we will use the continuous representation

$$I_{mn} = \int_{\bar{\omega}\in E_2} |\hat{\bar{\omega}}|^{m+n} \exp(i(m-n)\arg\hat{\bar{\omega}})|F(\bar{\omega})|^2\, d\bar{x}$$

$$= \int_{\bar{\omega}\in E_2} \hat{\bar{\omega}}^2 |F_1|^2\, d\bar{x},$$

where F_1 is a remapped version of the local Fourier spectrum, $F(\bar{\omega}) = \mathcal{F}(\phi(\bar{x})g(\bar{x}))(\bar{\omega})$, ϕ is the original image, and g is a window function. According to Theorem 2, the magnitudes of I_{mn} determine the minimum and maximum of $e(\bar{k})$ or, equivalently, they determine the minimum of $e_1(\bar{k})$ in analogy with Eq. (4):

$$\min_{\|\bar{k}\|=1} e_1(\bar{k}) = \min_{\|\bar{k}\|=1} \int_{\bar{\omega}\in E_2} (\bar{k}^t\bar{\omega})^2 |F_1(\bar{\omega})|^2\, d\bar{x}$$

$$= \min_{\theta} \int_{\bar{x}\in E_2} \left(\cos\theta \frac{\partial f_1}{\partial x} + \sin\theta \frac{\partial f_1}{\partial y}\right)^2 d\bar{x}. \tag{29}$$

Here f_1 is the inverse Fourier transform of F_1 and $\bar{k} = (\cos\theta, \sin\theta)^t$, and we have applied Parseval's theorem in order to pass to the spatial domain. The operator $\mathcal{L}$, defined as

$$\mathcal{L} = \varepsilon\cos\theta\frac{\partial}{\partial x} + \varepsilon\sin\theta\frac{\partial}{\partial y}, \tag{30}$$

is an infinitesimal generator of the two-parameter Lie group of transformations, the translation group [8]. The invariant curves of this group are constant lines with the directions determined only by θ since they satisfy $\mathcal{L}f_1 = 0$ for any $\varepsilon > 0$. Thus I_{mn} determines the closest member of this Lie group in the LSE sense. When $m - n = 2$, f_1 is the original local image f, and I_{mn} represents the direction along which a translation leaves the local image invariant (actually the direction yielding the smallest variance if translated). This can be seen easiest by inspecting Eq. (29) and noting that in the ideal case it equals zero. Thus the nonnegative integrand and thereby $\mathcal{L}f$ itself vanish. In the nonideal case this complex moment represents the LSE approximation of the solution lines of $\mathcal{L}f = 0$. For higher order symmetries, the exact analytical interpretation of I_{mn} in terms of invariants is analogous to the second-order symmetries. In this case I_{mn} represents the translation invariant direction of an $(m-n)/2$ "folded" local image. Because the folding is actually carried out in the local power spectrum this corresponds to equalizing the orientations differing with $2\pi/(m-n)$ in the spatial domain. Using infinitesimal operators (or, equivalently, local deformations caused by coordinate transformations) has been argued to be the most systematic way of analyzing a geometrical structure [2, 16, 23, 33]. One advantage of the complex moment approach relative to the Lie operator is its simplicity: it does not require more than ordinary algebra in order to derive the implementation. By contrast, the Lie algebra requires a thorough understanding of differential operators, but could result in more efficient implementations.

5.2. *Relation to Prolate Spheroidal Functions*

Another interesting connection to a known theory can be obtained when the complex moments of the local power spectrum are written in polar coordinates,

$$I_{mn} = \int_{\bar{\omega}\in E_2} r^{2n} r^{m-n} \exp(i(m-n)\varphi)|F(\bar{\omega})|^2\, d\bar{x}$$

$$= \int_{\bar{\omega}\in E_2} r^{2n}\hat{\bar{\omega}}^{m-n}|F(\bar{\omega})|^2\, d\bar{x}$$

$$= \int_{\bar{x}\in E_2} \left[\left(\frac{\partial}{\partial x} + i\frac{\partial}{\partial y}\right)^{(m-n)/2} f'(\bar{x})\right]^2 d\bar{x}, \tag{31}$$

where $r = |\hat{\bar{\omega}}|$, $\varphi = \arg\hat{\bar{\omega}}$, and f' is the inverse Fourier transform of r^nF. The distortion of the local spectrum through a weighting with radial polynomials can be compensated for by frequency weighting functions (see also [7]), or it can be seen as part of the discrete interpolation

function as will be explained below. Equation (31) makes it possible to compute the complex moment features directly in the spatial domain; see [4] for $m - n = 2$. Equation (31) can be approximated through smoothing the square of the complex-valued image

$$h(\bar{x}) = \sum_l f_l \left(\frac{\partial}{\partial x} + i\frac{\partial}{\partial y}\right)^{(m-n)/2} \mu_l(\bar{x}) \qquad (32)$$

on a discrete spatial checkerboard grid employing the theory of band-limited functions. Here μ_l is the interpolation function governing the behavior of the discrete image between the grid points and f_l is the original discrete image. In practice such an interpolation function is chosen to be a rotationally symmetric compact function (a function of $\|\bar{x}\|$). It can therefore be assumed that μ_l also contains the isotropic compensation for r^n. In that case h corresponds to the convolution result of the input image with a complex filter of the form

$$g_{mn}(|\hat{\bar{x}}|) \exp\left(i\frac{m-n}{2} \arg \hat{\bar{x}}\right), \qquad (33)$$

where g_{mn} is a one-dimensional function which, apart from m and n, also depends on the choice of the interpolation function. A special subset of these filters are prolate spheroidal functions [35]. The advantages of this type of filter in connection with rotation invariant pattern recognition have been argued in [12, 38], however, without referring to moments of the local power spectrum as discussed here. As for the Lie operators discussed above, operators like (33) can lead to more efficient implementations because the Gabor decomposition does not need to be incorporated in them. The operator set size (one spatial convolution per operator) instead of the Gabor filter set size (repeated filtering in the frequency domain) then determines the efficiency.

6. CONCLUSIONS

1. Complex moments provide for optimal solutions to the line-fitting, cross-fitting, etc., problems.

2. If computed on the basis of the local power spectrum, they yield estimates of the N-folded symmetry of the local image content.

3. If computed for all frequency bands individually, they allow for a simultaneous symmetry description at different scales. The moment set for each frequency band forms a complete representation.

4. The complex number representation is advantageous in circumventing the continuity, group representation, and range problems in pattern recognition.

5. Complex moments of the local power spectrum are related to the use of infinitesimal operators in the spatial domain, and therefore to the Lie algebra. They are also related to prolate spheroidal functions. Hence, they can be implemented by means of direct operations in the spatial domain. This could allow for a fast computation on special systems (e.g., pipeline or massively parallel architectures).

REFERENCES

1. Y. S. Abu-Mostafa and D. Psaltis, Recognitive aspects of moment invariants, *IEEE Trans. Pattern Anal. Machine Intell.* **PAMI-6,** 1984, 698–706.
2. J. Bigün, Pattern recognition by detection of local symmetries, in *Pattern Recognition and Artificial Intelligence* (E. S. Gelsema and L. N. Kanal, Eds.), pp. 75–90, North-Holland, The Netherlands, 1988.
3. J. Bigün, A structure feature for image processing applications based on spiral functions, *Comput. Vision Graphics Image Process.* **51,** 1990, 166–194.
4. J. Bigün, G. H. Granlund, and J. Wiklund, Multidimensional orientation estimation with application to texture analysis and optical flow, *IEEE Trans. Pattern Anal. Mach. Intell.* **PAMI-13,** 1991, 775–790.
5. J. Bigün, Frequency and orientation sensitive texture measures using linear symmetry, *Signal Process.* **29,** 1992, 1–16.
6. J. Bigün, Unsupervised feature reduction in image segmentation by local transforms, *Pattern Recogn. Lett.* **14,** 1993, 573–583.
7. J. Bigün and J. M. H. du Buf, *N*-folded symmetries by complex moments in Gabor space, *IEEE Trans. Pattern Anal. Mach. Intell.*, **16**(1), 1994, 80–87.
8. G. W. Bluman and S. Kumei, *Symmetries and Differential Equations*, Springer, Berlin, 1989.
9. P. Brodatz, *Textures: A Photographic Album for Artists and Designers*, Dover, New York, 1966.
10. J. M. H. du Buf and P. Heitkämper, Texture features based on Gabor phase, *Signal Process.* **23,** 1991, 227–244.
11. J. M. H. du Buf, Responses of simple cells: Events, interferences, and ambiguities, *Biol. Cybernet.* **68,** 1993, 321–333.
12. P. E. Danielsson, Rotation invariant linear operators with directional response, in *Proceedings 5th International Conference on Pattern Recognition, Miami Beach, FL, 1980,* pp. 1171–1176.
13. S. A. Dudani, K. J. Breeding, and R. McGhee, Aircraft identification by moment invariants, *IEEE Trans. Comput.* **C-26,** 1977, 39–45.
14. E. B. Elliott, *Algebra of Quantics,* 2nd ed., chap. 15, Oxford Univ. Press, New York, 1913.
15. M. Ferraro and T. M. Caelli, Relationship between integral transform invariances and Lie group theory, *J. Opt. Soc. Amer.* **A5,** 1988, 738–742.
16. L. M. J. Florack, B. M. T. H. Romeny, J. J. Koenderink, and M. A. Viergever, General intensity transformations and second order invariants, in *Proceed. 7th Scandinavian Conference on Image Analysis, Aalborg, Denmark, 1991,* pp. 338–345.
17. M. O. Freeman and B. E. A. Saleh, Moment invariants in the space and frequency domains, *J. Opt. Soc. Amer.* **5,** 1988, 1073–1084.
18. W. T. Freeman and E. H. Adelson, The design and use of steerable filters, *IEEE Trans. Pattern Anal. Mach. Intell.* **PAMI-13,** 1991, 891–906.
19. J. L. Gallant, J. Braun, and D. C. Van Essen, Selectivity for polar,

hyperbolic, and Cartesian gratings in macque visual cortex, *Science* **259,** 1993, 100–103.

20. G. H. Granlund, In search of a general picture processing operator, *Comput. Graphics Image Process.* **8,** 1978, 155–173.
21. W. Heisenberg, *The Physical Principles of the Quantum Theory,* Dover, New York, 1930.
22. W. C. Hoffman, The Lie algebra of visual perception, *J. Math. Psychol.* **3,** 1966, 65–98.
23. W. C. Hoffman, Higher visual perception as prolongation of the basic Lie transformation group, *Math. Biosci.* **6,** 1970, 437–471.
24. T. C. Hsia, A note on invariant moments in image processing, *IEEE Trans. Systems Man Cybernet.* **SMC-11,** 1981, 831–834.
25. M. K. Hu, Visual pattern recognition by moment invariants, *IRE Trans. Inform. Theory* **IT-8,** 1962, 179–187.
26. A. K. Jain and F. Farrokhnia, Unsupervised texture segmentation using Gabor filters, *Pattern Recognit.* **24,** 1991, 1167–1186.
27. A. Khotanzad and Y. H. Hong, Invariant image recognition by Zernike moments, *IEEE Trans. Pattern Anal. Mach. Intell.* **PAMI-12,** 1990, 489–497.
28. T. V. Papathomas and B. Julesz, Lie differential operators in animal and machine vision, in *From Pixels to Features* (J. C. Simon, Ed.), pp. 115–126, Elsevier, 1989.
29. P. Perona, Steerable-scalable kernels for edge detection and junction analysis, in *Proceedings 2nd European Conference on Computer Vision, Italy, 1992,* pp. 3–18.
30. M. Porat and Y. Y. Zeevi, The generalized Gabor scheme of image representation in biological and machine vision, *IEEE Trans. Pattern Anal. Mach. Intell.* **PAMI-10,** 1988, 452–468.
31. S. S. Reddi, Radial and angular invariants for image identification, *IEEE Trans. Pattern Anal. Mach. Intell.* **PAMI-3,** 1981, 240–242.
32. T. R. Reed and J. M. H. du Buf, A review of recent texture segmentation and feature extraction techniques, *Comput. Vision Graphics Image Process. Image Understanding* **57,** 1993, 359–372.
33. J. Segman, J. Rubinstein, and Y. Y. Zeevi, The canonical coordinates method for pattern deformation: Theoretical and computational considerations, *IEEE Trans. Pattern Anal. Mach. Intell.* **PAMI-14,** 1992, 1171–1183.
34. Y. Sheng, C. Lejeune, and H. H. Arsenault, Frequency-domain Fourier–Mellin descriptors for invariant pattern recognition, *Opt. Engrg.* **27,** 1988, 354–357.
35. D. Slepian, Prolate spheroidal wave functions, Fourier analysis, and uncertainty IV: Extensions to many dimensions; generalized prolate spheroidal functions, *Bell Systems Tech. J.* **43,** 1964, 3009–3057.
36. M. R. Teague, Image analysis via the general theory of moments, *J. Opt. Soc. Amer.* **70,** 1980, 920–930.
37. C. H. Teh and R. Chin, On image analysis by the methods of moments, *IEEE Trans. Pattern Anal. Mach. Intell.* **PAMI-10,** 1988, 496–513.
38. R. Wilson and M. Spann, Finite prolate spheroidal sequences and their application II: Image feature description and segmentation, *IEEE Trans. Pattern Anal. Mach. Intell.* **PAMI-10,** 1988, 193–203.
39. C. T. Zahn and R. Z. Roskiez, Fourier descriptors for plane closed curves, *IEEE Trans. Comput.* **C-21,** 1972, 269–281.
40. F. Zernike, Beugunstheorie des Schneidenverfahrens und seiner verbesserten Form der Phasenkontrastmethode, *Physica* **1,** 1934, 689–704.

JOSEF BIGÜN received his M.S. degree in 1983 and his Ph.D. degree in computer vision in 1988 from Linköping University, Sweden. He is now a senior research associate at the signal processing laboratory of the Swiss Federal Institute of Technology in Lausanne which he joined in 1988. His research interests include pattern recognition in general and texture and image sequence analysis in particular. He is currently conducting the research efforts of a joint project on motion analysis with Thomson-CSF, France. At the national as well as the international level, he participated in organizations of several conferences promoting pattern recognition and served as board or executive member of several scientific associations related to pattern recognition.

HANS DU BUF received the M.Eng. degree in electrical engineering from the Eindhoven University of Technology, the Netherlands in 1982. He gained his Ph.D. from the same university in 1987. His Ph.D work concerned visual psychophysics, in particular spatial brightness and contrast perception. He joined the Signal Processing Laboratory at the Swiss Federal Institute of Technology in Lausanne in 1987, where he worked on combining low-level image analysis and the modeling of visual perception by means of Gabor filters. He was appointed associate professor in signal and systems analysis at the University of Algarve (Portugal) in 1994.