REFERENCES

[1] S. Ullman, *The Interpretation of Visual Motion*. Cambridge, MA: MIT Press, 1979.

[2] H. C. Longuet-Higgins and K. Prazdny, "The interpretation of a moving retinal image," *Proc. R. Soc. London B*, vol. 208, pp. 385–397, 1980.

[3] K. Prazdny, "Egomotion and relative depth map from optical flow," *Biological Cybernetics*, vol. 36, pp. 87–102, 1980.

[4] ——, "Determining the instantaneous direction of motion from optical flow generated by a curvilinearly moving observer," *Computer Graphics and Image Processing*, vol. 17, pp. 238–248, 1981.

[5] A. R. Bruss and B. K. P. Horn, "Passive navigation," *Computer Vision, Graphics, and Image Processing*, vol. 21, pp. 3–20, 1983.

[6] J. H. Rieger and D. T. Lawton, "Processing differential image motion," *J. Opt. Soc. Am. A*, vol. 2, pp. 354–359, Feb. 1985.

[7] A. M. Waxman and K. Wohn, "Image flow theory: A framework for 3-D inference from time-varying imagery," in C. Brown, Ed., *Advances in Computer Vision, Volume 1*. New York: Lawrence Braum Associates, 1988.

[8] R. Y. Tsai and T. S. Wang, "Uniqueness and estimation of three-dimensional motion parameters of rigid objects with curved surfaces," *IEEE. Trans. Pattern Anal. Machine Intell.*, vol. PAMI-6, pp. 13–27, 1984.

[9] G. Adiv, "Determining three-dimensional motion and structure from optical flow generated by several moving objects," *IEEE. Trans. Pattern Anal. Machine Intell.*, vol. PAMI-7, no. 4, pp. 384–401, July 1985.

[10] A. Verri and T. Poggio, "Against quantitative optical flow," *Int. Conf. Computer Vision*, June 1987, pp. 171–180.

[11] G. Adiv, "Inherent ambiguities in recovering 3-D motion and structure from a noise flow field," *IEEE. Trans. Pattern Anal. Machine Intell.*, vol. PAMI-11, no. 5, pp. 477–489, May 1989.

[12] W. B. Thompson and J. K. Kearney, "Inexact vision," *Workshop on Motion, Representation, and Analysis*, pp. 15–22, May 1986.

[13] R. C. Nelson and J. Aloimonos, "Obstacle avoidance using flow field divergence," *IEEE. Trans. Pattern Anal. Machine Intell.*, vol. PAMI-11, no. 10, pp. 1102–1106, Oct. 1989.

[14] K. I. Beverley and D. Regan, "Texture changes versus size changes as stimuli for motion in depth," *Vision Res.*, vol. 23, no. 12, pp. 1387–1400, 1983.

[15] R. C. Nelson and J. Aloimonos, "Finding motion parameters from spherical motion fields (or the advantages of having eyes in the back of your head)," *Biological Cybernetics*, vol. 58, pp. 261–273, 1988.

[16] J. J. Koenderink and A. J. van Doorn, "Invariant properties of the motion parallax field due to the movement of rigid bodies relative to an observer," *Optica Acta*, vol. 22, no. 9, pp. 773–791, 1975.

[17] ——, "Local structure of movement parallax of the plane," *J. Opt. Soc. of Am. A*, vol. 66, no. 7, pp. 717–723, July 1976.

[18] J. Koenderink, "Optic flow," *Vision Res.*, vol. 26, no. 1, pp. 161–180, 1986.

[19] M. Subbarao, "Interpretation of visual motion: A computational study," Center for Automation Res., Univ. of Maryland, College Park, Tech. Rep. CAR-TR-221, Sept. 1986.

[20] ——, "Bounds on translational and angular velocity components from first order derivatives of image flow," *Proc. AAAI '87*, pp. 744–748.

[21] A. M. Goodbody, *Cartesian Tensors*. New York: Wiley, 1982.

[22] B. K. P. Horn and B. G. Schunck, "Determining optical flow," *Artificial Intelligence*, vol. 17, pp. 185–203, 1981.

[23] A. N. Tikhonov and V. Y. Arsenin, *Solutions of Ill-Posed Problems*. Wiley, 1977.

[24] D. L. Ringach, "A diffusion network for obstacle detection based on size-change information," M.S. thesis, Dept. of Elec. Eng., Technion, Israel Inst. of Tech., Haifa, Israel, 1990, in preparation (in Hebrew).

[25] H. Wagner, "Flow-field variables trigger landing in flies," *Nature*, vol. 297, pp. 147–148, May 1982.

[26] D. N. Lee, "A theory of visual control of braking based on information about time-to-collision," *Perception*, vol. 5, pp. 437–459, 1976.

[27] W. Schift and M. L. Detwiler, "Information used in judging impending collision," *Perception*, vol. 8, pp. 647–658, 1979.

[28] M. N. Davies and P. R. Green, "Head-bobbing during walking, running, and flying: Relative motion perception in the pigeon," *J. Exp. Biology*, vol. 138, pp. 71–91, 1988.

[29] S. Young and V. A. Taylor, "Visually guided chases in Polyphemus Pediculus," *J. Exp. Biology*, vol. 137, pp. 387–398, 1988.

[30] A. Friedman, *Generalized Functions and Partial Differential Equations*. Englewood Cliffs, NJ: Prentice-Hall Inc., 1963.

[31] ——, *Partial Differential Equations of Parabolic Type*. Englewood Cliffs, NJ: Prentice-Hall Inc., 1964.

[32] E. Zauderer, *Partial Differential Equations of Applied Mathematics, 2nd Edition*. New York: Wiley, 1989.

[33] J. M. Berezanskii, *Expansions in Eigenfunctions of Selfadjoint Operators*, vol. 17 of *Translations of Mathematical Monographs*. New York: American Mathematical Society, 1968.

[34] H. Weyl, "Über die asymptotische verteilung der eigenwerte," *Gott Nach.*, pp. 110–117, 1911.

[35] J. R. Kuttler and V. G. Sigillito, "Eigenvalues of the Laplacian in two dimensions," *SIAM Rev.*, vol. 26, no. 2, pp. 163–193, April 1984.

[36] M. H. Protter, "Can one hear the shape of a drum? Revisited," *SIAM Rev.*, vol. 29, no. 2, pp. 185–197, 1987.

[37] I. M. Gel'fand and G. E. Shilov, *Generalized Functions: Theory of Differential Equations*, vol. 3. New York: Academic Press Inc., 1964.

[38] H. F. Weinberger, "An isoperimetric inquality for the n-dimensional free membrane problem," *J. Rational Mech. Anal.*, vol. 5, pp. 533–636, 1956.

[39] L. E. Payne and H. F. Weinberger, "An optimal Poincare inequality for convex domains," *Arch. Rational Mech. Anal.*, vol. 5, pp. 286–292, 1960.

[40] J. H. Bramble and L. E. Payne, "Bounds in the Neumann problem for second order uniformly elliptic partial differential equations," *Pacific J. Math.*, vol. 12, pp. 823–833, 1962.

[41] D. W. Fox and W. C. Rheinboldt, "Computational methods for determining lower bounds for eigenvalues of operators in {Hilbert} space," *SIAM Rev.*, vol. 8, no. 4, pp. 427–462, Oct. 1966.

[42] V. G. Sigillito, *Explicit a-priori Inequalities with Applications to Boundary Value Problems*. New York: Pitman, 1977.

[43] R. C. Nelson and J. Aloimonos, "Using flow field divergence for obstacle avoidance in visual navigation," in *Proc. DARPA Image Understanding Workshop*, April 1988, pp. 548–567.

[44] G. E. Forsythe and W. R. Wasow, *Finite-Difference Equations for Partial Differential Equations*. New York: Wiley, 1960.

[45] F. B. Hildebrand, *Finite-Difference Equations and Simulations*. Englewood Cliffs, NJ: Prentice-Hall, 1968.

[46] A. R. Mitchell and D. F. Griffiths, *The Finite Difference Method in Partial Differential Equations*. New York: Wiley, 1980.

[47] A. Hurlbert and T. Poggio, "A network for image segmentation using color," in D. S. Touretzky, Ed., *Advances in Neural Information Processing Systems I*. New York: Morgan Kaufmann, 1989.

[48] T. Meis and U. Marcowitz, *Numerical Solution of Partial Differential Equations*. New York: Springer-Verlag, 1978.

[49] Y. Barniv, "Velocity filtering applied to optical flow calculations," NASA Tech. Mem. 102802, Aug. 1990.

[50] D. L. Ringach and Y. Baram, "Obstacle detection by diffusion," Computer Science Dept., Technion, Israel Institute of Technology, CIS Report 9220, Nov. 1992.

N-folded Symmetries by Complex Moments in Gabor Space and Their Application to Unsupervised Texture Segmentation

Josef Bigün and J. M. Hans du Buf

***Abstract*— Complex moments of the Gabor power spectrum yield estimates of the *N*-folded symmetry of the local image content at different frequency scales, that is, they allow to detect linear, rectangular,**

Manuscript received October 8, 1991; revised December 28, 1992. This work is partly supported by Thomson-CSF (France). Recommended for acceptance by Editor-in-Chief A. K. Jain.

J. Bigün and J. M. H. du Buf are with the Signal Processing Laboratory, Swiss Federal Institute of Technology, EPFL-Ecublens, CH-1015 Lausanne, Switzerland.

IEEE Log Number 9212237.

hexagonal/triangular, and so on, structures with very fine to very coarse resolutions. Results from experiments on the unsupervised segmentation of real textures indicate their importance for image processing applications. Real geometric moments computed in Gabor space also provide for very powerful texture features, but lack the clear geometrical interpretation of complex moments.

***Index Terms*—*N*-folded symmetries, real geometric moments, Legendre moments, complex moments, Zernike moments, Gabor spectral decomposition, texture features, unsupervised segmentation.**

I. Introduction

A local spectral decomposition by means of a bank of anisotropic filters allows to discriminate between different textures. Many studies have been performed, demonstrating in one way or another the significance of the local power spectrum [11], [20], [25], [28], [31] or phase spectrum [11], [14], [15]. However, the power or phase spectrum itself is exploited, rather than extracting syntactically or semantically meaningful primitives. Very few but semantically significant parameters which describe the shape of the local power spectrum can be quite powerful in texture discrimination. Du Buf [16] has shown that a set of parameters indicating whether a texture is isotropic or not, fine or coarse, rippled or rough, can be obtained by approximating the shape of the Gabor power spectrum with a Gaussian in a least-squares sense. Bigün [4], [5] has shown that the detection of many higher-order primitives can be modeled as an abstract line detection problem. A multiscale version of the simplest type of these primitives yields linear symmetry features. At each scale these provide for local orientation estimates along with certainties, therefore describing whether a texture is fine or coarse, isotropic or anisotropic [8].

Real (geometrical) and complex moments have often been used for object recognition. Hu [24] proposed invariants based on geometrical moments as features for the discrimination between objects with different shapes. Complex moments were proposed by Reddi [27] in order to facilitate the derivation of geometric invariants. Next, we will show that real and complex moments are both able to code the shape of the Gabor power spectrum. While real moments enable a "blind" mathematical description, and therefore a limited interpretability, complex moments yield geometrically significant image attributes. Complex moments are related to the line-fitting, cross-fitting, and so on problems, actually giving optimal solutions in the least-squares error sense. Therefore, complex moments when applied in Gabor space allow for an interpretation of the local image content in terms of N-folded symmetries. The symmetry order N refers to the complexity of the image structure, because $N/2$ gives the number of dominant local orientations: a two-folded symmetry means one orientation and therefore a linear structure, a four-folded symmetry means a rectangular structure with two orthogonal orientations, a 6-folded symmetry implies hexagonal structures with three orientations, and so on. We will use the terms linear, rectangular, and hexagonal in order to address such structures, but we emphasize that *any* structure with a N-folded symmetry is meant. Not only structures with rectangular and hexagonal but also square and triangular primitives, at fine as well as coarse resolution scales.

The rest of this correspondence is organised as follows. In Section II the Gabor spectral decomposition actually used in our experiments is presented. Geometric and Legendre moments of the Gabor power spectrum are introduced in Section III, and illustrated by a few experimental results on unsupervised texture segmentation. Complex moments of the Gabor power spectrum, together with experimental results, will be presented in Section IV. This correspondence will be concluded with a brief discussion in Section V.

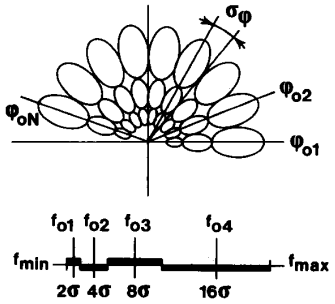


Fig. 1. The polar-separable filters give a rosette-like partition of the spectral half-plane (top), frequency bands doubling in width (bottom).

II. The Local Spectrum and the Gabor Decomposition

The complex spectrogram of an image is defined as the Fourier transform of the product of the image and the shifted version of a window function:

$$\mathcal{F}(\phi(\bar{x})g(\bar{x}-\bar{x}_0))(\bar{x}_0,\bar{\omega}_0) = \int_{\bar{x}\in E_2} \phi(\bar{x})g^*(\bar{x}-\bar{x}_0)\exp(-i\bar{\omega}_0^t\bar{x})dE_2, \quad (1)$$

where E_2 is the two dimensional Euclidean space, ϕ is the image, and g is the window function. As originally introduced by Gabor [22], the discretization of the local spectrum can be achieved by using a Gaussian

$$g(\bar{x}) = \frac{1}{2\pi\sigma^2}\exp(-\frac{\|\bar{x}\|^2}{2\sigma^2}) \quad (2)$$

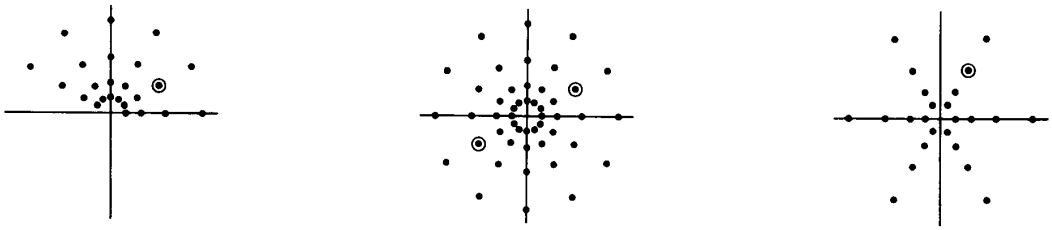
as the (2-D) window function. The local spectrum and the image are completely determined by the values of (1) on the Gabor lattice [22], i.e., $\bar{x}_0 \in X$ and $\bar{\omega}_0 \in \Omega$, with X and Ω being two "checker-board" lattices associated with the spatial and the frequency domains. Since we will perform the same operations at all positions $\bar{x}_0$, we will drop this variable or, equivalently, we will assume that $\bar{x}_0 = 0$. Moreover, only the squared modulus of the local spectrum

$$\rho(\bar{\omega}_0) = |\mathcal{F}(\phi(\bar{x})g(\bar{x}))(0,\bar{\omega}_0)|^2 \quad (3)$$

will concern us. This function will be referred to as the local power spectrum. Frequently we will use the notation $\rho(\omega_x,\omega_y)$ in Cartesian or $\rho(\omega_r,\omega_\varphi)$ in polar coordinates when meaning $\rho(\bar{\omega})$ for the sake of readability. Different choices of discretization lattices have been studied, e.g., [2], [18], [26]. However, for image analysis purposes the fixed size Gaussian is a handicap since it inevitably excludes harmonics with low frequencies. Another drawback is that it implicitly assumes a uniform frequency distribution, whereas real textures often have a linearly decreasing log power spectrum [19]. For these reasons, and in order to obtain isotropy in the orientation, we will choose our transfer functions variable in size and our discretization lattice as a rectangular checker-board in the log-polar frequency coordinate system.

Now, assuming polar-separable Gaussian transfer functions, the spectral half plane is partitioned into M frequency and N orientation bands:

$$G_{ij}(\omega_r,\omega_\varphi) = \exp\left\{\frac{-(\omega_r-\omega_{r_i}^0)^2}{2\sigma_{r_i}^2}\right\}\cdot\exp\left\{\frac{-(\omega_\varphi-\omega_{\varphi_j}^0)^2}{2\sigma_\varphi^2}\right\} \quad (4)$$



(a) (b) (c)

Fig. 2. Configurations used to compute real moments, from left to right the asymmetric, the symmetric, and the double-angle configuration.

with $1 \leq i \leq M$ and $1 \leq j \leq N$. The index pair i, j refers to a filter's position as well as its radial and angular bandwidth, whereas $\{\omega^0_{r_i}\}_i \otimes \{\omega^0_{\varphi_j}\}_j$ defines the discretization grid.

The application of polar-separable transfer functions results in a rosette-like partition of the frequency domain, see Fig. 1. The transfer functions are taken to overlap in such a way that neighbouring filters have equal values $1/\sqrt{e}$ at a distance σ_φ and σ_{r_i} from their corresponding grid points. The N orientations are taken equidistant, consequently $\sigma_\varphi = \pi/2N$. The angular bandwidth of the filters therefore equals π/N and the N filter orientations are

$$\omega^0_{\varphi_j} = 2\sigma_\varphi(j-1)\,;\; 1 \leq j \leq N. \tag{5}$$

The M frequency bands are approximately distributed in octave steps, such that the radial bandwidth doubles for each next higher-frequency band. Introducing $\omega_{\min}$ and $\omega_{\max}$ which position the inner and outer bounds of the rosette on the frequency scale (Fig. 1), it follows that $\sigma = (\omega_{\max} - \omega_{\min})/2(2^M - 1)$. Hence,

$$\left.\begin{array}{l} \omega^0_{r_i} = \omega_{\min} + \sigma\{1 + 3(2^{i-1} - 1)\} \\ \sigma_{r_i} = \sigma \cdot 2^{i-1} \end{array}\right\}\; 1 \leq i \leq M. \tag{6}$$

Applying this filter bank to an input image and computing the squared modulus of the complex filter results, we obtain an estimate of the local power spectrum for each position in the input image. The number of orientation and frequency channels is an important choice. Increasing the filter bank implies a better resolution of the local spectrum, and this will influence the frequency and orientation components which can be discriminated. We will use $M = 5$ and $N = 6$ as a reasonable compromise between CPU time and spectral resolution.

III. Real Moments of the Local Power Spectrum

In order to code the power spectrum with less parameters, one can also apply real moments [9]. These will be briefly reviewed here, and the results obtained with them are compared with those obtained with complex moments later on. Geometric moments of the discrete Gabor power spectrum are defined by

$$m_{pq} = \sum_{i,j} \omega^p_{x_i} \omega^q_{y_j} \rho(\omega_{x_i}, \omega_{y_j}) \tag{7}$$

with $p, q = 0, 1, 2, \cdots$. Introducing $\omega^c_x = m_{10}/m_{00}$ and $\omega^c_y = m_{01}/m_{00}$, we can compute the central moments

$$\mu_{pq} = \sum_{i,j} (\omega_{x_i} - \omega^c_x)^p (\omega_{y_j} - \omega^c_y)^q \rho(\omega_{x_i}, \omega_{y_j}). \tag{8}$$

Although real moments m_{pq} (or μ_{pq}) provide a complete representation (Weierstrass approximation theorem), the basis set $\omega^p_x \omega^q_y$ is not orthogonal. In addition, we will limit the order to $p + q \leq 4$ in our applications. Complete and orthogonal basis sets on the interval $[-1, 1]$ are provided by the Legendre polynomials $P_p(\omega_x)$ and $P_q(\omega_y)$, with

$$\int_{-1}^{1} P_n(\zeta) P_m(\zeta) d\zeta = \frac{2}{2m+1} \delta_{nm} \tag{9}$$

and

$$P_n(\zeta) = \sum_{\ell=0}^{n} a_{n\ell} \zeta^\ell = \frac{1}{2^n n!} \frac{d^n}{d\zeta^n} (\zeta^2 - 1)^n. \tag{10}$$

Legendre moments of the discrete local power spectrum are then defined as

$$\lambda_{pq} = \frac{(2p+1)(2q+1)}{4} \sum_{i,j} P_p(\omega_{x_i}) P_q(\omega_{y_j}) \rho(\omega_{x_i}, \omega_{y_j}), \tag{11}$$

in which ω_{x_i} and ω_{y_j} are scaled on the interval $[-1, 1]$. Further details on Legendre moments can be found in [30].

Application to Texture Segmentation

We use a bank of 30 filters, and compute real moments up to the fourth order. This means that we reduce the number of texture features in a first step from 30 to 15. From these 15 features we have to select a subset, the selected subset to be segmented by Spann and Wilson's [29] unsupervised quadtree-based segmentation algorithm. A *completely* unsupervised segmentation can be obtained by an automatic feature reduction utilizing either a *global* Karhunen–Loève transform (GKL) on the feature image set, or by means of a *local* Karhunen-Lo'eve transform (LKL). In the latter case the local scatter matrices are computed on Gaussian-weighted neighbourhoods, taking only the eigenvector with the maximum eigenvalue. From these local eigenvectors the global covariance matrix is then computed, and the resulting global eigenvectors are applied to the feature image set. The LKL transform has been shown to perform better than the GKL transform [6]. In all experiments we took the transformed feature images which correspond to trace-normalized eigenvalues greater than a prescribed threshold value (we used $\lambda \geq 0.01$). We note that the same Gabor decomposition, LKL transform, and quadtree segmentation were applied in all experiments performed; this fixed procedure allows for a comparison of the different feature extraction methods.

In the experiments, we considered both central and Legendre moments, and varied the frequency coordinates but only in computing the moment sets. This means that the Gabor spectral decomposition was always performed by applying the fixed filter set with five frequencies and six orientations as described above and depicted in Fig. 1, but that moments were computed on the basis of manipulated frequency coordinates as illustrated in Fig. 2. In other words, $\{\rho\}$ is considered as a set of values which can be associated with points in a plane in different ways. The reason for doing so is that moments are strongly influenced by the geometric configuration or a deformation thereof, sometimes improving texture discriminability. The three

configurations applied in the present study were: the asymmetric filter coordinates (see Fig. 2(a)), the symmetric coordinates (Fig. 2(b)), and the double-angle coordinates (Fig. 2(c)). The latter coordinate system is inspired by the fact that the filters with $j = 1$ and $j = N$ are neighbours [23], and corresponds to a special case of computing complex moments; see further on. The three configurations discussed above were also tried by using equidistant radial frequencies, that is $(\log \omega_{r_i}, \omega_{\varphi_j})$, giving six possibilities in total.

All six possibilities have been considered in computing central moments and Legendre moments, segmenting several test images by applying the LKL transform and the quadtree segmentation method. Figs. 3 and 4 show best results obtained with two test images, each containing 7 textures from the Brodatz album [12] and from aerial imagery, respectively. These seven textures are combined in a sixteen-patch mosaic in such a way that all combinations share a common boundary at least once. It should be mentioned that the image in Fig. 4 is extremely difficult to segment by using other texture feature extraction methods available in our lab, see [3], [13], giving very poor results in general. Best results, as shown in Figs. 3 and 4, have been obtained by central moments using the asymmetric configuration (Fig. 2(a)). The other configurations, equidistant radial frequencies, as well as Legendre moments gave worse results, that is four to six classes detected in general. In the case of Fig. 3, all regions could be detected and the boundaries found are very accurate. In the case of Fig. 4, two texture classes have been lost. In conclusion, excellent to very acceptable results can be obtained by computing central moments of the local power spectrum. However, the drawback of using real moments is that they yield a "blind" mathematical representation of the local power spectrum. In contrast, complex moments of the local power spectrum, to be described below, provide for a geometrically interpretable description of the local image content.

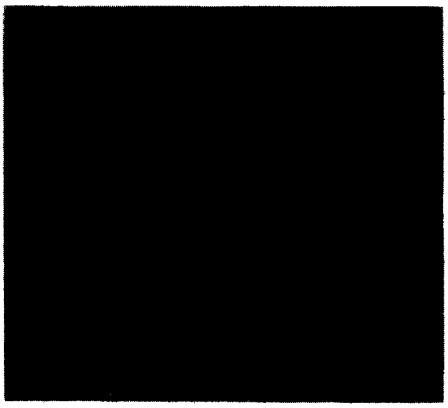
IV. Complex Moments and Symmetries

A. Brief Review

Complex moments arise naturally in connection with rotation invariant 2-D real functions. A version of these based on a set of radial and orthogonal polynomials, also known as Zernike moments, were introduced by Zernike [32] in 1934. In pattern recognition they appear much later although rotation invariant optical character recognition has been dealt with by Hu [24] already in 1962. The latter author proposed a set of rotation invariant features, known as moment invariants, based on geometrical (real valued) moments. Reddi [27] introduced a set of functions, similar to the complex moments which will be exploited here, simplifying considerably the analytic derivation of geometric moment invariants. He used the terms radial and angular moments for this set of functions, and showed that geometric moment invariants can be obtained by manipulating complex moments in a simple way, such as computing their magnitudes. These complex moments are defined by

$$I_{mn} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x+iy)^m (x-iy)^n \rho(x,y)dxdy, \qquad m,n = 0,1,2,\cdots. \tag{12}$$

A related approach to the derivation of geometric moment invariants using complex moments, and a visual interpretation of these, was proposed by Abu–Mostafa and Psaltis [1]. Their interpretation is based on the Fourier expansions of the spatial image along concentric circles. Although useful for object boundary recognition, this interpretation is not suitable for the local image characterization problem addressed here, since this expansion is essentially origin dependent and global, rather than translation invariant and local. Freeman and Saleh [21] proposed the use of spatial and frequency domain complex moment magnitudes (Fourier-Mellin descriptors as they refer to them) as features. Although they considered these in the context of object recognition, this approach is more in line with the ideas presented here, because it stresses the importance of features based on both domains. Next, we will present a novel interpretation of complex moments. This interpretation completes (in the mathematical sense) the linear symmetry description of a neighbourhood [8], and introduces N-folded symmetries. We will start with a reformulation of the line-fitting process, and generalize it to the cross-fitting process and beyond.



(a)

(b)

Fig. 3. (a) Test image with seven Brodatz textures in 16 patches, and (b) an unsupervised segmentation result obtained with 5 feature images which were computed by applying the LKL tranform to central moments. The boundaries in the original image are shown dashed for comparison.

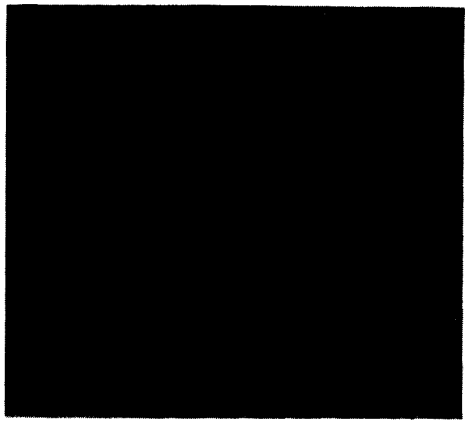
B. The Optimal Line- and Cross-Fitting Processes

Assume a two dimensional, real, and non-negative function $\rho(\bar{x})$, $\bar{x} \in E_2$. We would like to fit a line to this continuous function in the least-squares error sense (LSE). Fig. 5 (right) illustrates this problem:

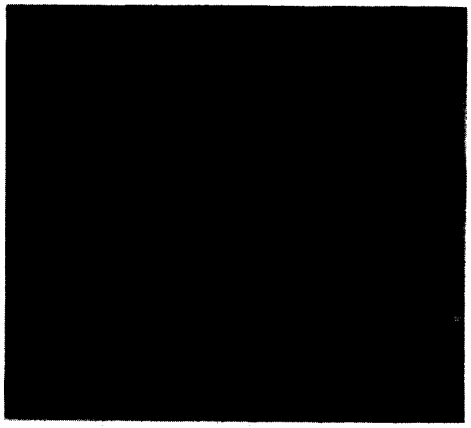
$$\min_{\|\bar{k}\|=1} e(\bar{k}) = \int_{\bar{x}\in E_2} \|\bar{x} - (\bar{k}^t\bar{x})\bar{k}\|^2 \rho(\bar{x})dE_2, \tag{13}$$

where $\|\bar{x} - (\bar{k}^t\bar{x})\bar{k}\|$ is the perpendicular distance of a point $\bar{x}$ to a line represented by the unit vector $\bar{k}$. This quadratic problem can be solved by using straightforward differential calculus. An alternative, the proof of which goes beyond the scope of this correspondence, is to use the complex moments (12), see [10]. For convenience we use the "bar-hat" notation: Let $\bar{x} \in E_2$, that is, $\bar{x}$ is a vector with two real components:

$$\bar{x} = (x_1, x_2)^t. \tag{14}$$



(a)



(b)

Fig. 4. (a) Test image with seven textures from aerial imagery, and (b) an unsupervised segmentation result obtained with central moments.

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

$$\hat{\bar{x}} = x_1 + ix_2. \tag{15}$$

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

$$\bar{\hat{x}} = (\mathrm{Re}(\hat{x}), \mathrm{Im}(\hat{x}))^t. \tag{16}$$

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 solution of (13) results then in 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

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

and the minimum and maximum error lines are determined by

$$\hat{\bar{k}}^2_{\min} = \frac{I_{20}}{|I_{20}|} \quad \text{and} \quad \hat{\bar{k}}^2_{\max} = -\frac{I_{20}}{|I_{20}|}. \tag{17}$$

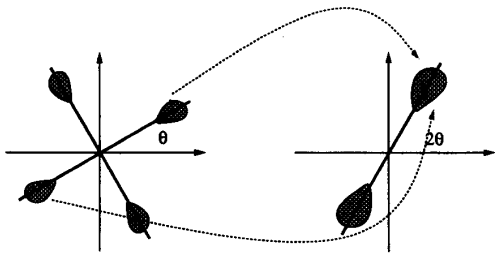


Fig. 5. The cross-fitting problem is solved by applying a coordinate transform first, followed by a line-fitting in the new coordinate system.

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 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. The transformation

$$\varphi_1 = 2\varphi \tag{18}$$

is two to one and maps a cross to a line, see Fig. 5. Hence, Theorem 1 can be applied to the transformed function

$$\rho_1(\bar{x}) = |\hat{\bar{x}}|^2 \left(\rho(|\hat{\bar{x}}|, \frac{\arg \hat{\bar{x}}}{2}) + \rho(|\hat{\bar{x}}|, \frac{\arg \hat{\bar{x}}}{2} + \pi) \right) \tag{19}$$

in order to fit a cross to the original function ρ. It can be shown that this transformation is not explicitly required because the result follows directly from the complex moment $I_{40}(\rho)$. A straightforward extension of the cross-fitting process to N-folded symmetry fitting, see [10], gives us the following general theorem.

Theorem 2: The complex moments $I_{mn}(\rho)$ and $I_{\frac{m+n}{2},\frac{m+n}{2}}(\rho)$, with $m - n \neq 0$, determine 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 $\frac{2\pi}{m-n}$ rotations. The error function of the fitting process is

$$\min_{\|\bar{k}_1\|=1} e(\bar{k}) = \int_{\bar{x} \in E_2} \|\bar{x} - (\bar{k}_1^t \bar{x})\bar{k}_1\|^2 \rho_1(\bar{x}) dE_2$$

where

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

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 error are given by

$$|I_{mn}| = e(\bar{k}_{\max}) - e(\bar{k}_{\min}) \tag{20}$$
$$I_{\frac{m+n}{2},\frac{m+n}{2}}(\rho) = e(\bar{k}_{\max}) + e(\bar{k}_{\min}), \tag{21}$$

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

$$\hat{\bar{k}}^{m-n}_{\min} = \frac{I_{m-n}}{|I_{m-n}|} \quad \text{and} \quad \hat{\bar{k}}^{m-n}_{\max} = -\hat{\bar{k}}^{m-n}_{\min}. \tag{22}$$

Although introduced in the general 2-D domain, Theorem 2 can be applied to the local power spectrum, yielding a somewhat different geometrical interpretation in the spatial domain. This is because the line-fitting process, for instance, in the local power spectrum corresponds to the detection of *any* structures with one dominant orientation, i.e., irrespective of the local phase spectrum.

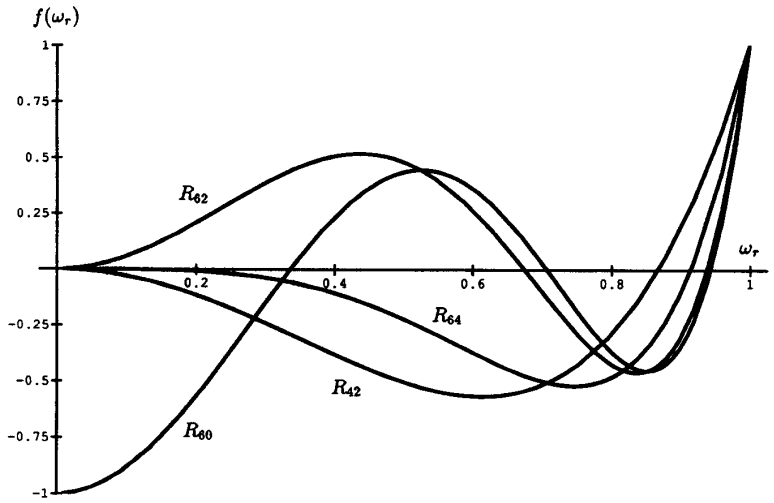


Fig. 6. Frequency weighting by Zernike polynomials.

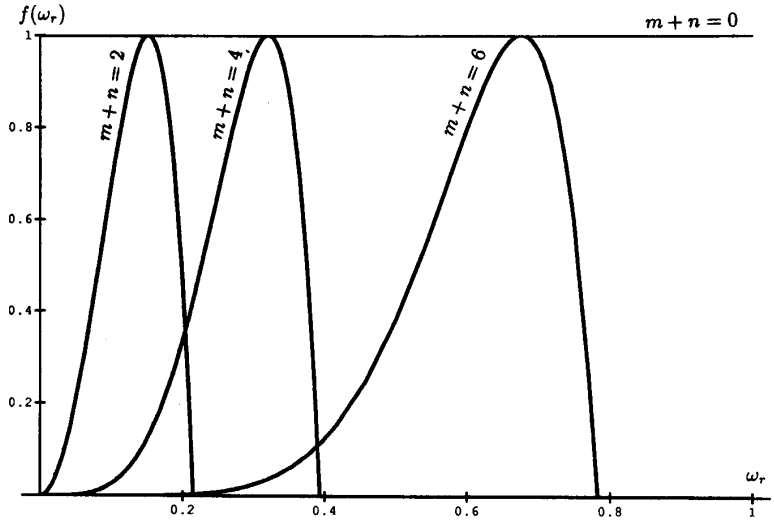


Fig. 7. Frequency weighting by octave bands.

One of the main reasons for exploiting the local power spectrum in texture discrimination is its translation invariance within homogeneous regions. The concentration of the power to a line reveals the existance of a texture with one distinct orientation (linear or two-folded symmetry) around an inspected point. The concentration to a cross means a texture composed of patterns with two, mutually orthogonal, directions and therefore rectangular patterns with a four-folded symmetry. Similarly, the existence of six-folded symmetry means textures with 3 orientations (hexagonal patterns), and so on. The arguments of the complex moments give the orientation of the estimated N-folded symmetry, while the magnitudes give measures of the estimation quality, that is certainties.

C. Frequency Distribution Models and Experimental Results

When a complex moment is computed by integration over the entire local power spectrum, we observe that the inherent minimization process distorts the original function ρ radially before the minimization actually takes place. The reason for why one might want to compute complex moments on the basis of the entire local spectrum, despite the distortion effect, is that otherwise one would obtain a high dimensional feature space. This would impose severe computational problems for succeeding clustering methods. The integration over the entire spectrum is one way to avoid computing every symmetry order for every frequency band. The inherent distortion is caused by the radial weighting of the original function ρ with a factor r^{m+n} (for notational convenience we will use r and φ instead of ω_r and ω_φ in the sequel). This can be seen by rewriting the complex moments (12) in the polar form:

$$I_{mn} = \int_0^{\infty} \int_0^{2\pi} f(r) \exp(i(m-n)\varphi)\rho(r,\varphi) r dr d\varphi \qquad (23)$$

where

$$f(r) = r^{m+n}. \qquad (24)$$

According to the Weierstrass theorem, moments and thereby complex moments are complete in $\mathcal{L}^2(C_u)$, C_u being the unit circle. But the

respective moments are still correlated. This problem can be circumvented by utilizing radial polynomials which make the complex moments orthogonal, such as those proposed by Zernike [30], other polynomials, or even other functions than polynomials which are less correlated. In our segmentation experiments we computed complex moments of orders 2, 4, and 6, and compared the following three frequency weighting methods:

1) *The natural weighting:* This model corresponds to when $f(r)$ in (23) is chosen as r^{m+n}. The experimental results learned that the associated features discriminate aerial as well as Brodatz textures poorer if compared to the octave model to be presented below. The comparison was done on the basis of the number of classes detected and the quality of the class boundaries.
2) *Zernike polynomials:* In this model, we choose $f(r)$ as Zernike polynomials, and the complex moments computed with this weighting are often referred to as Zernike moments. These polynomials are now given and illustrated in Fig. 6:

$$f(r) = R_{m+n,m-n}(r) = \sum_{s=0}^{m} (-1)^s \frac{(m+n-s)!}{s!(m-s)!(n-s)!} r^{(m+n-2s)}. \quad (25)$$

Here, m, n, and $m-n$ are assumed to be non-negative. In terms of segmentation results, this model performed equally well if compared to the natural weighting.
3) *Octave band weighting:* Inspired by the Laplacian pyramid [17] we pairwise subtract the complex moments of subsequent orders, which results in the following weighting function:

$$g_{m+n}(r) = a_{m+n} r^{m+n} - b_{m+n} r^{m+n+2}$$
$$f_{m+n}(r) = \begin{cases} g_{m+n}(r), & \text{if } g_{m+n}(r) > 0; \\ 0, & \text{otherwise.} \end{cases} \quad (26)$$

The scaling constants a_{m+n} and b_{m+n} are needed to force the maximum of f_{m+n} to be 1 and to occur at the exact octave filter positions. Figure 7 shows the frequency distribution for this model. Since each moment order is uniquely adapted to a specific frequency band, the complex moment order corresponds directly to the frequency order, i.e. a high $m+n$ indicates fine structures while a low one points towards coarse structures. This model is the only one among the three that we studied which attaches a geometric meaning both to the moment order $m+n$ and the symmetry order $m-n$ at the same time, thereby keeping the number of feature images relatively low.

The experimental results obtained with octave band weighting were far better than those obtained with the natural weighting and Zernike polynomials. A segmentation obtained by complex moments with the octave band weighting is shown in Fig. 8 . Six classes have been found. Fig. 9 shows the result when the octave model is applied to structured textures with two-, four-, and six-folded symmetries. In this case all seven classes were found.

V. Discussion

We have seen that real geometric moments of the local power spectrum provide for powerful texture features suitable for an unsupervised image segmentation. Best results were achieved by computing central moments on the basis of the asymmetrical frequency coordinates (Fig. 2(a)). Results obtained with the symmetric and double-angle representations (Figs. 2(b) and 2(c)), as well as these representations with equidistant radial frequencies, were clearly inferior. This effect can be explained by the strong asymmetry in the representation, which leads to moments with an increased discriminability. However, this effect also depends on the local power spectra and therefore on the textures present in an image, and cannot be generalized.

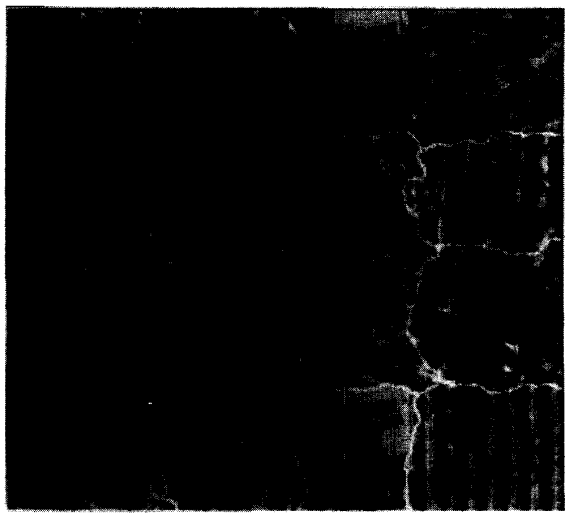


Fig. 8. Unsupervised segmentation obtained by octave-band weighted complex moments.

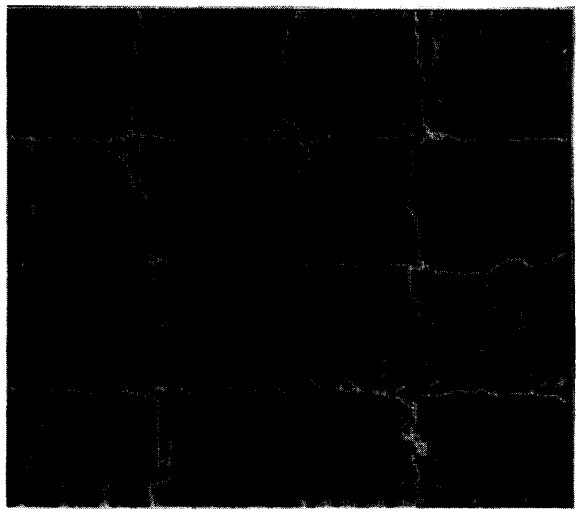


Fig. 9. Segmentation result by applying octave-band weighted moments to structured textures with two-, four-, and six-folded symmetries.

Despite the good results, real moments lack a clear interpretation of geometric structures. Complex moments allow for such an interpretation because they provide for optimal solutions to the line-fitting, cross-fitting, and so on, problems. When complex moments are computed with an octave-band weighting of the Gabor power spectrum, they enable a geometrical description at different frequency scales. Complex moments will perform best for textures with N-folded symmetries, in accordance with their geometrical significance. However, for stochastic textures such as in Fig. 8 their performance is less obvious because of the absent symmetries. The good segmentation result shown in Fig. 8 therefore indicates that complex moments, as for real moments, can provide for a complete representation of the local power spectrum, or an adequate representation if the moment order is limited.

There are a few practical problems which need further attention. The boundary artifacts as illustrated in Figs. 4 and 8, but which were more or less prominent in all experimental results, are caused

by the fact that certain (if not all) moments, real or complex, are sensitive to the mixture of two local power spectra at a texture boundary. This leads to deviant feature values at the boundaries. Hence, the prominence of this effect also depends on the segmentation method applied. We used Spann and Wilson's [29] quadtree based segmentation algorithm throughout our study, which allows for a comparison of the different moment models. This method is what is called region-based. The boundary refinement subprocess is clearly disturbed by the deviant feature values at the boundaries. An alternative, as discussed in [16], would be to combine region- and boundary-based methods.

Furthermore, the maximum moment order has to be somehow fixed or determined. This is not a trivial problem, because a large feature dimensionality leads to practical problems in the feature reduction and segmentation processes. Since we apply complex moments to the discrete Gabor power spectrum, the order depends on the number of orientations in the Gabor decomposition. This, in turn, depends on the pattern discriminability required. The very good results obtained by applying complex moments of order 6 to the spectrum of rather stochastic textures (Fig. 8) shows that isotropic textures can be dealt with. For structured textures, the symmetries present require an appropriate moment order and therefore a corresponding number of orientations in the Gabor filtering. However, this reasoning is not quite realistic because of the overlap of the modulation transfer functions of the filters. A spectral component which lies between two filters will result in a less peaked but broader contribution to the local power spectrum. In addition, the larger the symmetry order, the smaller the differences in orientations. Perhaps a number of six or eight orientations, leading to complex moments of order six or eight, is already sufficient for most applications.

REFERENCES

[1] Y. S. Abu-Mostafa and D. Psaltis, "Recognitive aspects of moment invariants," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-6, pp. 698-706, 1984.

[2] M. J. Bastiaans, "A sampling theorem for the complex spectrogram, and Gabor's expansion of a signal in Gaussian elementary signals," *Optical Eng.*, vol. 20, pp. 594-598, 1981.

[3] J. Bigün and G. H. Granlund, "Optimal orientation detection of linear symmetry," *First Int. Conf. on Computer Vision*, London, June 1987, pp. 433–438. IEEE Comp. Soc. Press.

[4] J. Bigün, "Pattern recognition by detection of local symmetries," in *Pattern recognition and artificial intelligence*, E. S. Gelsema and L. N. Kanal, Eds. Amsterdam, The Netherlands: North-Holland, 1988, pp. 75–90.

[5] ——, "A structure feature for image processing applications based on spiral functions," *Comput. Vision, Graphics Image Processing*, vol. 51, pp. 166–194, 1990.

[6] ——, "Unsupervised feature reduction in image segmentation by local transforms," *Pattern Recognit. Lett.*, vol. 14, pp. 573–583, 1993.

[7] J. Bigün, G. H. Granlund, and J. Wiklund, "Multidimensional orientation estimation with application to texture analysis and optical flow," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 13, pp. 775–790, 1991.

[8] J. Bigün, "Frequency and orientation sensitive texture measures using linear symmetry," *Signal Processing*, vol. 29, pp. 1–16, 1992.

[9] J. Bigün and J. M. H. du Buf, "Texture segmentation by real and complex moments of the Gabor power spectrum," in *Progress in Image Analysis and Processing*, V. Cantoni *et al.*, Eds. Singapore: World Scientific, 1992, pp. 191–198.

[10] ——, "Symmetry interpretation of complex moments and the local power spectrum," *Visual Commun. Image Representation*, under review.

[11] A. C. Bovik, M. Clark, and W. S. Geisler, "Multichannel texture analysis using localized spatial filters," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 12, pp. 55–73, 1990.

[12] P. Brodatz, *Textures: A Photographic Album for Artists and Designers*. New York: Dover, 1966.

[13] J. M. H. du Buf, M. Kardan, and M. Spann, "Texture feature performance for image segmentation," *Pattern Recognit.*, vol. 23, pp. 291–309, 1990.

[14] J. M. H. du Buf, "Gabor phase in texture discrimination," *Signal Processing*, vol. 21, pp. 221–240, 1990.

[15] J. M. H. du Buf and P. Heitkämper, "Texture features based on Gabor phase," *Signal Processing*, vol. 23, pp. 227–244, 1991.

[16] J. M. H. du Buf, "Abstract processes in texture discrimination," *Spatial Vision*, vol. 6, pp. 221–242, 1992.

[17] P. Burt, "Fast filter transforms for image processing," *Computer Graphics and Image Processing*, vol. 16, pp. 20–51, 1981.

[18] I. Daubechies, "The wavelet transform, time-frequency localization and signal analysis," *IEEE Trans. Inform. Theory*, vol. 36, pp. 961–1005, 1990.

[19] D. J. Field, "Relations between the statistics of natural images and the response properties of cortical cells," *J. Opt. Soc. Amer.*, vol. A4, pp. 2379–2394, 1987.

[20] I. Fogel and D. Sagi, "Gabor filters as texture discriminator," *Biological. Cybern.*, vol. 61, pp. 103–113, 1989.

[21] M. O. Freeman and B. E. A. Saleh, "Moment invariants in the space and frequency domains," *J. Opt. Soc. Amer.*, vol. 5, pp. 1073–1084, 1988.

[22] D. Gabor, "Theory of communication," *J. Inst. Elect. Eng.*, vol. 93, pp. 429–459, 1946.

[23] G. H. Granlund, "In search of a general picture processing operator," *Comput. Graphics Image Processing*, vol. 8, pp. 155–173, 1978.

[24] M. K. Hu, "Visual pattern recognition by moment invariants," *IRE Trans. Inform. Theory*, vol. IT-8, pp. 179–187, 1962.

[25] A. K. Jain and F. Farrokhnia, "Unsupervised texture segmentation using Gabor filters," in *Proc. Int. Conf. Syst., Man, Cybern.*, Los Angeles, CA, Nov. 1990, pp. 14–19.

[26] M. Porat and Y. Y. Zeevi, "The generalized Gabor scheme of image representation in biological and machine vision," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 10, pp. 452–468, 1988.

[27] S. S. Reddi, "Radial and angular invariants for image identification," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-3, pp. 240–242, 1981.

[28] T. R. Reed and H. Wechsler, "Spatial-frequency representations for image segmentation and grouping," *Image and Vision Computing*, vol. 9, pp. 175–193, 1991.

[29] M. Spann and R. Wilson, "A quad-tree approach to image segmentation which combines statistical and spatial information," *Pattern Recognit.*, vol. 18, pp. 257–269, 1985.

[30] C. H. Teh and R. Chin, "On image analysis by the methods of moments," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 10, pp. 496–513, 1988.

[31] R. Wilson and M. Spann, "Finite prolate spheroidal sequences and their application II: Image feature description and segmentation," *IEEE Trans. Pattern Anal. and Machine Intell.*, vol. 10, pp. 193–203, 1988.

[32] F. Zernike, "Beugungstheorie des Schneidenverfahrens und seiner verbesserten Form der Phasenkontrastmethode," *Physica*, vol. 1, pp. 689–704, 1934.