Rotational Invariance Based on Fourier Analysis in Polar and Spherical Coordinates

Qing Wang, Olaf Ronneberger, and
Hans Burkhardt, *Member, IEEE*

Abstract—In this paper, polar and spherical Fourier analysis are defined as the decomposition of a function in terms of eigenfunctions of the Laplacian with the eigenfunctions being separable in the corresponding coordinates. The proposed transforms provide effective decompositions of an image into basic patterns with simple radial and angular structures. The theory is compactly presented with an emphasis on the analogy to the normal Fourier transform. The relation between the polar or spherical Fourier transform and the normal Fourier transform is explored. As examples of applications, rotation-invariant descriptors based on polar and spherical Fourier coefficients are tested on pattern classification problems.

Index Terms—Invariants, Fourier analysis, radial transform, multidimensional.

---◆---

1 INTRODUCTION

NOTHING needs to be said about the importance of Fourier transform in image processing and pattern recognition. Usually, Fourier transform is formulated in Cartesian coordinates, where a separable basis function in 3D space without normalization is given by

$$e^{i\mathbf{k}\cdot\mathbf{r}} = e^{ik_x x}e^{ik_y y}e^{ik_z z}, \qquad (1)$$

where (x, y, z) are coordinates of the position $\mathbf{r}$ and (k_x, k_y, k_z) are components of the wave vector $\mathbf{k}$ along the corresponding axis. The basis function (1) represents a plane wave, which is a periodic pattern. Fourier analysis of an image is, therefore, the decomposition of the image into the basic patterns represented by (1).

The Laplacian is an important operator in mathematics and physics. Its eigenvalue problem gives the time-independent wave equation. In Cartesian coordinates, the operator is written as

$$\nabla^2 = \nabla_x^2 + \nabla_y^2 + \nabla_z^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}.$$

Equation (1) is an eigenfunction of the Laplacian and is separable in Cartesian coordinates.

When defined on the whole space, functions given in (1) are mutually orthogonal for different $\mathbf{k}$; wave vectors take continuous values and it is said that one has a continuous spectrum. Over finite regions, the mutual orthogonality generally does not hold. To get an orthogonal basis, $\mathbf{k}$ can only take values from a discrete set and the spectrum becomes discrete. The continuous Fourier transform reduces to Fourier series or to the discrete Fourier transform.

For objects with certain rotational symmetry, it is more effective for them to be investigated in polar (2D) or spherical (3D)

● *The authors are with the Computer Science Department and the Centre for Biological Signalling Studies, University of Freiburg, Chair of Pattern Recognition and Image Processing, Georges-Koehler-Allee Geb. 052, D-79110 Freiburg, Germany.*
E-mail: {qwang, ronneber, hans.burkhardt}@informatik.uni-freiburg.de.

coordinates. It would be of great advantage if the image can be decomposed into wave-like basic patterns that have clear radial and angular structures. Ideally, this decomposition should be an extension of the normal Fourier analysis and can, therefore, be called Fourier analysis in the corresponding coordinates. To fulfill these requirements, the basis functions should take the separation-of-variable form

$$R(r)\Phi(\varphi) \qquad (2)$$

for 2D and

$$R(r)\Theta(\vartheta)\Phi(\varphi) = R(r)\Omega(\vartheta, \varphi) \qquad (3)$$

for 3D, where (r, φ) and (r, ϑ, φ) are the polar and spherical coordinates, respectively. They should also be the eigenfunctions of the Laplacian so that they represent wave-like patterns and that the associated transform is closely related to the normal Fourier transform. The concrete form of the angular and radial parts of the basis functions will be investigated and elaborated in the coming sections but will be briefly introduced below in order to show previous work related to them.

For polar coordinates, the angular part of a basis function is given by

$$\Phi(\varphi) = \frac{1}{\sqrt{2\pi}}e^{im\varphi}, \qquad (4)$$

where m is an integer. The associated transform in angular coordinate is just the normal 1D Fourier transform. For spherical coordinates, the angular part is a spherical harmonic

$$\Omega(\vartheta, \varphi) = Y_{lm}(\vartheta, \varphi) = \sqrt{\frac{2l+1}{4\pi}\frac{(l-m)!}{(l+m)!}}P_{lm}(\cos\vartheta)e^{im\varphi}, \qquad (5)$$

where P_{lm} is an associated Legendre polynomial and l and m are integers, $l \geq 0$ and $|m| \leq l$. The corresponding transform is called Spherical Harmonic (SH) transform and has been widely used in representation and registration of 3D shapes [6], [7], [8].

The angular parts of the transforms in 2D and 3D are, therefore, very familiar. Not so well known are the transforms in the radial direction. The radial basis function is a Bessel function $J_m(kr)$ for polar coordinates and a spherical Bessel function $j_l(kr)$ for spherical coordinates. In both cases, the parameter k can take either continuous or discrete values, depending on whether the region is infinite or finite. For functions defined on $(0, \infty)$, the transform with $J_m(kr)$ as integral kernel and r as weight is known as the Hankel transform. For functions defined on a finite interval, with zero-value boundary condition for the basis functions, one gets the Fourier-Bessel series [1]. Although the theory on Fourier-Bessel series has long been available, it, mainly, has applications in physics-related areas [16], [17]. Zana and Cesar's work [10] and a few references therein are the only ones we can find that employ Fourier-Bessel series expansion for 2D image analysis. Methods based on Zernike moments are, on the other hand, much more popular in applications where we believe the Fourier-Bessel expansion also fits.

The SH transform works on the spherical surface. To describe 3D volume data, one can use a collection of SH features calculated on concentric spherical surfaces of different radii, as suggested in [7]. This approach treats each spherical surface as independent to one another; therefore, it cannot describe the radial structures effectively. This observation motivated the whole work presented here.

In this paper, the operations that transform a function into the coefficients of the basis functions, given in (2) and (3) and described above, will simply be called *polar* and *spherical Fourier transform*, respectively. It should be noted though that, in the literature, the former often refers to the normal Fourier transform

with wave vectors expressed in polar coordinates [14] and the latter often refers to the SH transform [15].

Due to the extreme importance of the Laplacian in physics, the expansion of functions with respect to its eigenfunctions is naturally not new there (e.g., [18], [19]). The idea that these eigenfunctions can be used as basis functions for analyzing 2D or 3D images is unfamiliar to the pattern recognition society. There also lacks a simple and systematic presentation of the expansion from the point of view of signal analysis. Therefore, although part of the derivation is scattered in books like [1], we rederive the basis functions to emphasize the analogy to the normal Fourier transform. Employment of the Sturm-Liouville theory makes this analogy clearer and the derivation more compact.

The proposed polar and spherical Fourier transforms are connected with the normal Fourier transform by the Laplacian. We investigate the relation between them so that one can understand the proposed transforms from another point of view. It is found that the relation also provides computational convenience. An advantage of the polar and the spherical Fourier transforms is that rotation-invariant descriptors can be very easily extracted from the transform coefficients. We will show how to do this.

Section 2 deals with the polar Fourier transform. Besides presentation of the theory, issues about calculation of the coefficients are discussed. A short comparison between polar Fourier basis functions and Zernike functions is made at the end. Parallel to Section 2, the theory for the spherical Fourier transform is summarized in Section 3. As examples for the application of the theory, rotation-invariant descriptors based on the polar and spherical Fourier coefficients are applied to object classification tasks in Section 4. At the end, conclusion and outlook are given.

To avoid the ideas being hidden by too much mathematics, the derivation and theories are only outlined in this paper. Those who are interested in more details are referred to [23].

2 POLAR FOURIER TRANSFORM

2.1 Basis Functions

2.1.1 Helmholtz Equation and Angular Basis Functions

As an extension from the Cartesian case, we begin with the eigenfunctions of the Laplacian, whose expression in polar coordinates is given by

$$\nabla^2 = \nabla_r^2 + \frac{1}{r^2}\nabla_\varphi^2, \tag{6}$$

where

$$\nabla_r^2 = \frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial}{\partial r}\right) \quad \text{and} \quad \nabla_\varphi^2 = \frac{\partial^2}{\partial \varphi^2} \tag{7}$$

are the radial and angular parts. The eigenvalue problem can be written as

$$\nabla_r^2\Psi(r,\varphi) + \frac{1}{r^2}\nabla_\varphi^2\Psi(r,\varphi) + k^2\Psi(r,\varphi) = 0, \tag{8}$$

which is the Helmholtz equation in polar coordinates. Only nonnegative k^2 are of interest. We further require $k \geq 0$. By substituting the separable form $\Psi(r,\varphi) = R(r)\Phi(\varphi)$ into (8), one gets

$$\frac{\partial^2}{\partial \varphi^2}\Phi + m^2\Phi = 0, \tag{9}$$

$$\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial}{\partial r}\right)R + \left(k^2 - \frac{m^2}{r^2}\right)R = 0. \tag{10}$$

The solution to (9) is simply

$$\Phi_m(\varphi) = \frac{1}{\sqrt{2\pi}}e^{im\varphi}, \tag{11}$$

with m being an integer.

2.1.2 Radial Basis Functions

The general nonsingular solution to (10) is

$$R(r) = J_m(kr), \tag{12}$$

where J_m is the mth order Bessel function. Bessel functions satisfy the orthogonality relation

$$\int_0^\infty J_m(k_1 r)J_m(k_2 r)r\,dr = \frac{1}{k_1}\delta(k_1 - k_2), \tag{13}$$

just like the complex exponential functions satisfy

$$\int_{-\infty}^\infty e^{ik_1 x}\left[e^{ik_2 x}\right]^* dx = 2\pi\delta(k_1 - k_2). \tag{14}$$

Actually, $J_m(kr)$ with $k \geq 0$ forms an orthogonal basis for functions defined on $(0,\infty)$.

For the normal Fourier transform, an infinite space corresponds to a continuous spectrum and a finite space to a discrete spectrum, which is selected by proper boundary conditions. The same also applies to the radial basis functions (12). Over the finite interval $(0,a)$, the orthogonality relation, like in (13), generally does not hold any more, instead,

$$\int_0^a J_m(k_1 r)J_m(k_2 r)r\,dr$$
$$= \frac{a}{k_1^2 - k_2^2}\left[k_2 J_m(k_1 a)J'_m(k_2 a) - k_1 J_m(k_2 a)J'_m(k_1 a)\right]. \tag{15}$$

Boundary conditions need to be imposed on $R(r) = J_m(kr)$ to select a set of k values so that the corresponding functions are mutually orthogonal. According to the Sturm-Liouville (S-L) theory [2], for all the nonnegative k satisfying the boundary condition

$$R(a)\cos\beta - aR'(a)\sin\beta = 0, \qquad \beta \in [0,\pi), \tag{16}$$

and leaving $R(r) = J_m(kr)$ as nonzero functions, the eigenfunctions $\{R(r) = J_m(kr)\}$ form an orthogonal basis on $(0,a)$.

With $x = ka$ and $R(r) = J_m(kr)$, (16) becomes

$$J_m(x)\cos\beta - xJ'_m(x)\sin\beta = 0. \tag{17}$$

Suppose $(x_{m1} < x_{m2} < \cdots < x_{mn} < \cdots)$ are the nonnegative solutions to (17) and $J_m(x_{mn}r/a)$ are nonzero functions,

$$k_{nm} = \frac{x_{mn}}{a}, \qquad n = 1, 2, \ldots \tag{18}$$

(the indices n and m now exchange their order for the sake of convention) are then the solutions to (16). The corresponding eigenfunctions are orthogonal to each other:

$$\int_0^a J_m(k_{nm}r)J_m(k_{n'm}r)r\,dr = N_n^{(m)}\delta_{nn'}, \tag{19}$$

where $N_n^{(m)}$ can be determined from (10) and (15) as

$$N_n^{(m)} = \frac{a^2}{2}\left[J'_m{}^2(x_{mn}) + \left(1 - \frac{m^2}{x_{mn}^2}\right)J_m^2(x_{mn})\right]. \tag{20}$$

The normalized radial function can therefore be defined as

$$R_{nm}(r) = \frac{1}{\sqrt{N_n^{(m)}}}J_m(k_{nm}r). \tag{21}$$

The set of functions $\{R_{nm}|n = 1, 2, \ldots\}$ forms an orthonormal basis with weight r on $(0, a)$. A function $f(r)$ defined on this interval can be expanded as

$$f(r) = \sum_{n=1}^{\infty} \left[\int_0^a f(\rho) R_{nm}(\rho) \rho \, d\rho \right] R_{nm}(r). \tag{22}$$

In (17), β can take any values in $[0, \pi)$. Two cases are especially interesting:

Zero-value boundary condition. With $\sin \beta = 0$, (17) reduces to

$$J_m(x) = 0 \tag{23}$$

(note that $x = ka$). x_{mn} should be the positive zeros of $J_m(x)$. Under this condition,

$$N_n^{(m)} = \frac{a^2}{2} J_{m+1}^2(x_{mn}), \tag{24}$$

and the right-hand side of (22) is usually known as the mth order Fourier-Bessel series of $f(r)$.

Derivative boundary condition. With $\cos \beta = 0$, (17) becomes

$$J_m'(x) = 0. \tag{25}$$

x_{mn} should be the positive zeros of $J_m'(x)$ except for one special case: $x_{01} = 0$. This case is not explicitly covered in [1]. One can verify that $x = 0$ is a solution to $J_0'(x) = 0$ and $J_0(0 \cdot r/a) = 1$ is nonzero and is, indeed, an eigenfunction of the S-L system. Under the derivative boundary condition, we have

$$N_n^{(m)} = \frac{a^2}{2} \left(1 - \frac{m^2}{x_{mn}^2} \right) J_m^2(x_{mn}), \tag{26}$$

with the special case $N_1^{(0)} = a^2/2$.

Different boundary conditions lead to different spectra of the system. The choice should depend on the problems under investigation. To give an impression how the radial functions look like, we show the first few of them for $m = 2$ with zero and derivative boundary conditions in Figs. 1a and 1b. It is intuitive to choose the zero boundary condition when the images tend to be zero at $r = a$ and the derivative condition when the images tend to be constant in radial direction near $r = a$. Often, it is necessary to do some experiments to find the better choice.

The asymptotic behavior of the Bessel functions [1] reveals the wave-like property of the radial function: for $k_{nm}r \gg |m^2 - \frac{1}{4}|$,

$$R_{nm}(r) \sim \frac{1}{\sqrt{r}} \cos \left(k_{nm}r - \frac{m\pi}{2} - \frac{\pi}{4} \right). \tag{27}$$

$R_{nm}(r)$ approaches a cosine function with the amplitude decreasing as fast as $1/\sqrt{r}$. There is a phase shift of $-(m\pi/2 + \pi/4)$, which corresponds to a "delay" of the function to take the wave-like form near the origin (see Fig. 6a).

2.1.3 Basis Functions

The basis function for the polar Fourier transform comprises the radial and the angular parts. Consequently, for the transform defined on the whole space, the basis function is given by

$$\Psi_{k,m}(r, \varphi) = \sqrt{k} J_m(kr) \Phi_m(\varphi) \tag{28}$$

with k taking continuous nonnegative values. For the transform defined on the finite region $r < a$, the basis function is given by

$$\Psi_{nm}(r, \varphi) = R_{nm}(r) \Phi_m(\varphi). \tag{29}$$

It satisfies the equation

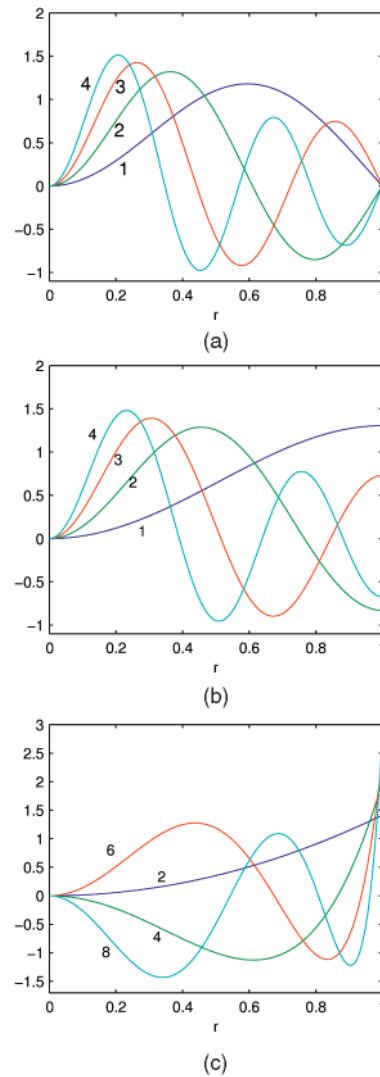$$\nabla^2 \Psi_{nm} + k_{nm}^2 \Psi_{nm} = 0. \tag{30}$$

Fig. 1. The first few radial basis functions for 2D with $m = 2$ and $a = 1$: (a) R_{nm} with zero boundary condition, (b) R_{nm} with derivative boundary condition, and (c) normalized radial Zernike function $\tilde{Z}_{nm}$. The number beside each curve is the value of n.

Those familiar with quantum mechanics will recognize k_{nm}^2 as the energy level (except for a constant factor) of the system. The basis functions $\Psi_{nm}(r, \varphi)$ with the lowest energy levels are shown in Fig. 2. One can find that the higher the energy level, the finer the structures. Therefore, for image analysis, the value of k is an indication of the scale of the basic patterns, which is consistent with the normal Fourier transform.

2.2 Expansion and Rotation-Invariant Descriptors

A 2D function $f(r, \varphi)$ defined on the whole space can be expanded with respect to $\Psi_{k,m}$, as defined in (28):

$$f(r, \varphi) = \int_0^{\infty} \sum_{m=-\infty}^{\infty} P_{k,m} \Psi_{k,m}(r, \varphi) k \, dk, \tag{31}$$

where

$$P_{k,m} = \int_0^{\infty} \int_0^{2\pi} f(r, \varphi) \Psi_{k,m}^*(r, \varphi) r \, dr \, d\varphi \tag{32}$$

are the polar Fourier coefficients (P stands for $Polar$). The infinite transform, as given in (31) and (32), is mainly of theoretical interest. In practice, one should use the transform defined on a
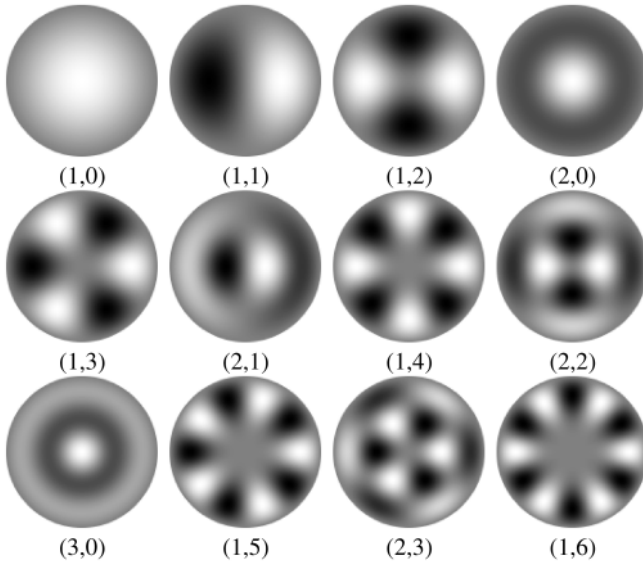
Fig. 2. Basic patterns represented by Ψ_{nm} with zero boundary condition. Shown are the real parts of the functions. (n, m) pairs are given under each pattern. The patterns are listed in the increasing order of the value of k_{nm}.

finite region. A function $f(r, \varphi)$, defined on $r < a$, can be expanded with respect to $\{\Psi_{nm}\}$ as

$$f(r, \varphi) = \sum_{n=1}^{\infty} \sum_{m=-\infty}^{\infty} P_{nm} \Psi_{nm}(r, \varphi), \qquad (33)$$

where the coefficients

$$P_{nm} = \int_0^a \int_0^{2\pi} f(r, \varphi) \Psi_{nm}^*(r, \varphi) r \, dr \, d\varphi. \qquad (34)$$

There are two indices for the expansion. How should the terms be ordered, and therefore, be truncated for a finite-term expansion? A natural way is according to the energy levels. In the language of image analysis, according to the scales of the basic patterns. Larger scale patterns should be taken into account first. This is often the best choice if no other information about the data is available. Fig. 3 shows the isolines of k_{nm}.

The polar Fourier coefficients provide a complete representation of the original function. But it is often desirable to have a rotation-invariant description of an object. A complete set of descriptors can be obtained by properly normalizing the coefficients according to the degree of rotational symmetry (similar to the technique in [21]). However, though phase information is very important, there still lacks a systematic and robust way of incorporating this information into the descriptors. For many problems, the modulus of the coefficients $|P_{nm}|$ make up a robust, although mathematically not complete, set of rotation-invariant descriptors. In this paper, $|P_{nm}|$ will be called Polar Fourier Descriptors (PFD). For real images, $|P_{n(-m)}| = |P_{nm}|$, one needs only to consider descriptors with nonnegative m.

A digital image is usually given on an equally-spaced grid in Cartesian coordinates. To evaluate the coefficients as in (34), it is advisable to map the image I into polar coordinates, where the transform becomes separable and the angular part can be done fast with FFT. The radial part then only needs to be performed for each kept term. The grid density of the mapped image I_{polar} should be high enough to accommodate the finest patterns in the expansion. Let the largest values for m and k_{nm} be $m_{\max}$ and $k_{\max}$. Denote the radial and angular size of I_{polar} as M_r and M_φ. The sampling theorem requires $M_\varphi \geq 2m_{\max}$. M_φ should also be chosen to facilitate fast calculations. Our numerical experiments
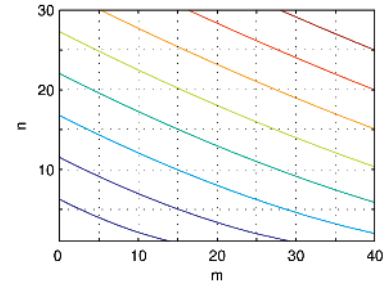

Fig. 3. Isolines of k_{nm}.

on the radial basis function show that it is enough to have $M_r \geq ak_{\max}/\pi + 25$ to ensure that (19) holds with a relative error about or less than 1 percent. Mapping of the image is a process of interpolating and sampling. It must be handled carefully to avoid aliasing. If the resolution in I_{polar} is coarser than in the original image I, one can either first smooth I then perform the mapping, or alternatively, first map I to polar coordinates with proper resolutions followed by smoothing and downscaling in r or φ. Which approach to take depends on the aspect ratio of I_{polar}.

2.3 Relation to Normal Fourier Transform in 2D

2.3.1 Infinite Transform

To find the relation between the polar and the normal Fourier transforms, one needs to know the relation of their bases. With the Jacobi-Anger Identity [4], the basis function for the normal Fourier transform can be expanded as

$$\frac{1}{2\pi} e^{i\mathbf{k} \cdot \mathbf{r}} = \frac{1}{2\pi} e^{ikr\cos(\varphi - \varphi_k)}$$
$$= \sum_{m=-\infty}^{\infty} \frac{i^m}{\sqrt{2\pi k}} e^{-im\varphi_k} \Psi_{k,m}(r, \varphi), \qquad (35)$$

where $\mathbf{k}$ is the wave vector and (k, φ_k) and (r, φ) are the polar coordinates of $\mathbf{k}$ and $\mathbf{r}$, respectively. $\Psi_{k,m}$ is defined in (28).

Suppose a function $f(r, \varphi)$ is defined on the whole space and its normal Fourier transform is C_{k,φ_k} (C stands for Cartesian, k and φ_k are written as subscripts for consistency of notation although they take continuous values), then it can be expressed as

$$f(r, \varphi) = \int_0^{\infty} \int_0^{2\pi} C_{k,\varphi_k} \left[\frac{1}{2\pi} e^{ikr\cos(\varphi - \varphi_k)} \right] k \, dk \, d\varphi_k. \qquad (36)$$

Substituting (35) into (36), one has the expansion of $f(r, \varphi)$ in $\Psi_{k,m}$ and can get the polar Fourier coefficient $P_{k,m}$ expressed in C_{k,φ_k} as

$$P_{k,m} = \frac{i^m}{\sqrt{k}} \frac{1}{\sqrt{2\pi}} \int_0^{2\pi} C_{k,\varphi_k} e^{-im\varphi_k} d\varphi_k. \qquad (37)$$

The relation is very simple. Except for the factor $(i^m/\sqrt{k})$, $P_{k,m}$ is just the Fourier coefficient of C_{k,φ_k} by considering φ_k as the variable (see Fig. 4a for an illustration).

2.3.2 Transform on Finite Regions

Strictly speaking, it is ambiguous to talk about the relation between the polar and the normal Fourier transforms when they are defined on finite regions because the shapes of the regions are different. We consider such a situation here: The normal Fourier transform is defined on a rectangle that is centered at the origin and encloses the disk where the polar Fourier transform is defined. Let the area of the rectangle be A.
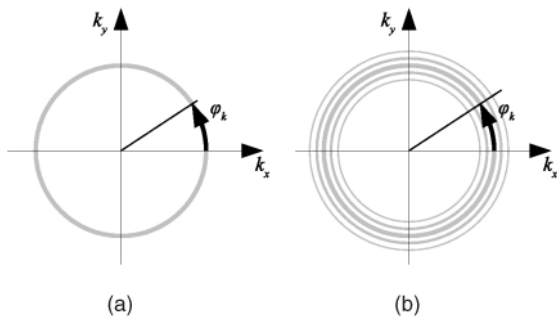
Fig. 4. Illustration of the relation of the polar Fourier coefficients and the normal Fourier coefficients C_{k,φ_k}. (a) When the space is infinite, $P_{k,m}$ is the Fourier coefficient of C_{k,φ_k}, with φ_k as the variable. (b) When the space is finite, P_{nm} is the weighted sum of the Fourier coefficients on different circles. With zero boundary condition, the weight function is proportional to (44).

Fig. 5. $\left|\frac{k_0 J_m(k_0 a)}{k_0^2 - k_{nm}^2}\right|$ as a function of k_0 for $n = 8, m = 5$, and $a = 32$, where $k_{nm} = 0.994$.

We first expand a plane wave in $\{\Psi_{nm}\}$ on the disk. As shown in (35), a plane wave can be expanded in $\Psi_{k,m}$, which, in turn, can be expanded easily in $\{\Psi_{nm}\}$ on the disk. We have

$$e^{i\mathbf{k}\cdot\mathbf{r}} = \sum_{n,m} i^m \sqrt{2\pi} \left[\int_0^a R_{nm}(\rho) J_m(k\rho)\rho \, d\rho \right] \cdot e^{-im\varphi_k} \Psi_{nm}(r,\varphi) \tag{38}$$

for $r < a$. The expression inside the square brackets is the coefficient of $J_m(kr)$ in $R_{nm}(r)$. It can be calculated by making use of (15). If k_{nm} are selected with the zero boundary condition, then

$$e^{i\mathbf{k}\cdot\mathbf{r}} = \sum_{n,m} (-1)^n i^m 2\sqrt{\pi} k_{nm} \frac{J_m(ka)}{k^2 - k_{nm}^2} e^{-im\varphi_k} \Psi_{nm}(r,\varphi). \tag{39}$$

This equation holds for any $\mathbf{k}$, including naturally those appearing in the normal Fourier transform defined on the rectangle, which we denote as $\mathbf{k}_0$.

A function $f(r,\varphi)$ defined on the disk can be extended to the rectangle by padding. Let the normal Fourier coefficients for the padded function be $C_{\mathbf{k}_0}$. On the disk,

$$f(r,\varphi) = \sum_{\mathbf{k}_0} C_{\mathbf{k}_0} \frac{1}{\sqrt{A}} e^{i\mathbf{k}_0 \cdot \mathbf{r}}. \tag{40}$$

The function $f(r,\varphi)$ can as well be expanded in $\{\Psi_{nm}\}$. Let the coefficient for Ψ_{nm} be P_{nm}. With the expansion (39), one can get

$$P_{nm} = (-1)^n i^m \frac{2\sqrt{\pi}}{\sqrt{A}} k_{nm} \sum_{\mathbf{k}_0} \frac{J_m(k_0 a)}{k_0^2 - k_{nm}^2} e^{-im\varphi_{k_0}} C_{\mathbf{k}_0} \tag{41}$$

for the zero boundary condition. We rewrite the main parts of $P_{k,m}$ and P_{nm} from (37) and (41) for comparison,

$$P_{k,m} \sim \int d\mathbf{k}' \, \delta(k' - k) \, e^{-im\varphi_{k'}} C_{k',\varphi_{k'}}, \tag{42}$$

$$P_{nm} \sim \sum_{\mathbf{k}_0} \frac{J_m(k_0 a)}{k_0^2 - k_{nm}^2} e^{-im\varphi_{k_0}} C_{\mathbf{k}_0}. \tag{43}$$

When the space becomes finite, the integral over the wave vector is replaced by a summation and the sharp function of the wavenumber $\delta(k' - k)$ is replaced by a more spreading one (see Fig. 4b for an illustration)
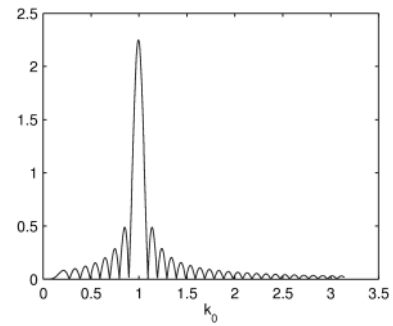
$$\frac{J_m(k_0 a)}{k_0^2 - k_{nm}^2}, \tag{44}$$

which has its maximum absolute value at $k_0 = k_{nm}$. According to the asymptotic behavior of Bessel functions, it arrives to its first zeros approximately at $|k_0 - k_{nm}| = \pi/a$ and will oscillatingly decrease on both sides. Fig. 5 shows the absolute value of (44) multiplied by k_0, which comes from the fact that the number of pixels at radius k is approximately proportional to k.

For completeness, if $\{\Psi_{nm}\}$ is determined with the derivative boundary condition, we can get

$$P_{nm} = (-1)^n i^m \frac{2\sqrt{\pi} a k_{nm}}{\sqrt{A}\sqrt{k_{nm}^2 a^2 - m^2}} \sum_{\mathbf{k}_0} \frac{k J_m'(ka)}{k^2 - k_{nm}^2} e^{-im\phi_{k_0}} C_{\mathbf{k}_0}. \tag{45}$$

Equations (41) and (45) can be used to calculate the polar Fourier coefficients P_{nm} from the normal Fourier coefficients $C_{\mathbf{k}_0}$, which can be obtained by FFT. This approach implies sinc interpolation in the spatial domain and is best suited when the underlying original signal is band-limited.

2.4   Comparison with Zernike Polynomials

Besides the basis functions defined with (29), there exists an infinity of sets of basis functions on a disk. One of the most famous are Zernike polynomials. Since Teh and Chin [9] made a comparison study on different moment methods, which shows that Zernike moments outperform other moment-based methods in terms of overall performance, there are a lot of applications using Zernike moments, e.g., [11], [12], [13]. Zernike functions are defined on a unit disk, and, when expressed in polar coordinates, have the following form [3]:

$$V_{nm}(r,\varphi) = Z_{nm}(r) e^{im\varphi}, \tag{46}$$

where m is any integer, $n \geq 0$ is an integer and is the order of the polynomial, $n \geq |m|$, and $n - |m|$ is even. The angular part is the same as that of (29). The radial Zernike function Z_{nm} is a polynomial in r:

$$Z_{nm}(r) = \sum_{s=0}^{\frac{n-|m|}{2}} (-1)^s \frac{(n-s)!}{s!\left(\frac{n+|m|}{2} - s\right)!\left(\frac{n-|m|}{2} - s\right)!} r^{n-2s}. \tag{47}$$

It has $(n - |m|)/2$ zeros between 0 and 1. The orthogonality relation of the radial functions is given by

$$\int_0^1 Z_{nm}(r) Z_{n'm}(r) r \, dr = \frac{1}{2n+2} \delta_{nn'}. \tag{48}$$

For comparison, we define the normalized radial function as

$$\widetilde{Z}_{nm} = \sqrt{2n+2} Z_{nm}. \tag{49}$$

The first few normalized radial functions for $m = 2$ are shown in Fig. 1c. The typical form of $\widetilde{Z}_{nm}$ with relatively large m and n is shown in Fig. 6 together with R_{nm} for comparison. One can find
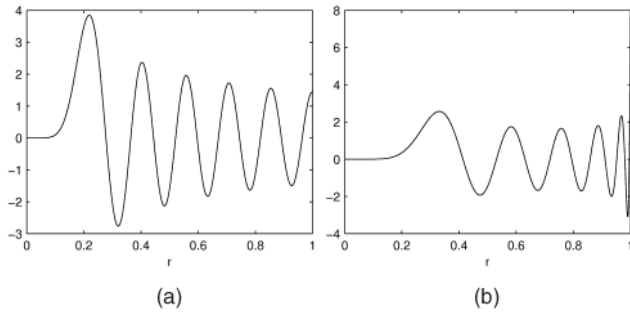
Fig. 6. (a) $R_{(11)8}$ as defined in (21) with derivative boundary condition for $a = 1$ and (b) $\widetilde{Z}_{(28)8}$ as defined in (49). Both have 10 zeros on $(0, 1)$.

that $\widetilde{Z}_{nm}$ has also a wave-like form. Like R_{nm}, it has also a "delay" near the origin for the "wave" to begin. Unlike R_{nm}, the amplitude of the "wave" does not decrease monotonically; instead, the "wavelength" decreases with r.

3   SPHERICAL FOURIER TRANSFORM

The theory on the spherical Fourier transform is only summarized here. Interested readers are referred to [23] for more details.

The angular part of the basis function is a spherical harmonic $Y_{lm}(\vartheta, \varphi)$, as defined in (5). The corresponding radial part is

$$R(r) \sim j_l(kr), \tag{50}$$

where j_l is the so-called spherical Bessel function of order l and is related to the ordinary Bessel function by

$$j_l(x) = \sqrt{\frac{\pi}{2x}} J_{l+\frac{1}{2}}(x). \tag{51}$$

The normalized basis function for spherical Fourier transform defined on the whole space is

$$\Psi_{k,l,m}(r, \vartheta, \varphi) = \sqrt{\frac{2}{\pi}} k j_l(kr) Y_{lm}(\vartheta, \varphi), \tag{52}$$

where k takes continuous values. When defined on a solid sphere of radius a, the normalized basis function is

$$\Psi_{nlm}(r, \vartheta, \varphi) = \frac{1}{\sqrt{N_n^{(l)}}} j_l(k_{nl} r) Y_{lm}(\vartheta, \varphi), \tag{53}$$

where n is a positive integer; k_{nl} and $N_n^{(l)}$ are determined from the S-L boundary conditions. With the **zero-value boundary condition**, let $(x_{l1} < x_{l2} < \cdots < x_{ln} < \cdots)$ be the positive solutions to $j_l(x) = 0$, then $k_{nl} = x_{ln}/a$ and

$$N_n^{(l)} = \frac{a^3}{2} j_{l+1}^2(x_{ln}). \tag{54}$$

With the **derivative boundary condition**, let $(x_{l1}, x_{l2}, \ldots, x_{ln}, \ldots)$ be the increasingly ordered positive zeros of $j_l'(x)$ (except for $x_{01} = 0$), then $k_{nl} = x_{ln}/a$ and

$$N_n^{(l)} = \frac{a^3}{2} \left[1 - \frac{l(l+1)}{x_{ln}^2} \right] j_l^2(x_{ln}), \tag{55}$$

with the special case $N_1^{(0)} = a^3/3$.

A function $f(r, \vartheta, \varphi)$ that is defined on a solid sphere with radius a can be expanded in terms of $\Psi_{nlm}(r, \vartheta, \varphi)$:

$$f(r, \vartheta, \varphi) = \sum_{n=1}^{\infty} \sum_{l=0}^{\infty} \sum_{m=-l}^{l} S_{nlm} \Psi_{nlm}(r, \vartheta, \varphi), \tag{56}$$

where

$$S_{nlm} = \int_0^a \int_0^\pi \int_0^{2\pi} f(r, \vartheta, \varphi) \Psi_{nlm}^*(r, \vartheta, \varphi) r^2 \sin\vartheta \, dr \, d\vartheta \, d\varphi \tag{57}$$

are the spherical Fourier coefficients (S stands for Spherical).

Extracting rotational invariants of objects in 3D is slightly more complicated than in 2D as there are two angular coordinates now. It is well known that Y_{lm} with $-l \le m \le l$ span a subspace $\mathcal{Y}_l$ that is invariant with respect to the rotation group. For a function defined on the spherical surface and having expansion

$$f(\vartheta, \varphi) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} h_{lm} Y_{lm}(\vartheta, \varphi), \tag{58}$$

the magnitude of its projection onto $\mathcal{Y}_l$, i.e., $\sqrt{\sum_{m=-l}^{l} |h_{lm}|^2}$ remains unchanged under rotation and is a rotational invariant (will be called a SH descriptor) of the function. In [7], it is suggested to use SH descriptors on spherical surfaces of a set of radii to describe volume data. These descriptors will be called radius-wise SH Descriptors (rwSHD) in this paper. It is worth noting that a rwSHD should be scaled with the square of its radius to account for the change of the area of the spherical surface with the radius. It is straightforward to know that, for a function defined on a solid sphere, its projection onto the subspace spanned by $\Psi_{nlm} = R_{nl} Y_{lm}$ with $-l \le m \le l$ has also a fixed magnitude under rotation as all the variance of the basis function Ψ_{nlm} under rotation is captured by its angular part. That is to say,

$$\sqrt{\sum_{m=-l}^{l} |S_{nlm}|^2} \tag{59}$$

for any n and l is a rotation-invariant property of that function. We call (59) a Spherical Fourier Descriptor (SFD).

The relation between the spherical Fourier transform and the normal Fourier transform can be derived similarly as in the 2D case. Any function $f(r, \vartheta, \varphi)$ defined on the whole space can be expanded in either plane waves $(\frac{1}{\sqrt{2\pi}})^3 e^{i\mathbf{k} \cdot \mathbf{r}}$ or in spherical waves $\Psi_{k,l,m}$:

$$\begin{aligned} &f(r, \vartheta, \varphi) \\ &= \int_0^\infty \int_0^\pi \int_0^{2\pi} C_{k,\vartheta_k,\varphi_k} \left(\frac{1}{\sqrt{2\pi}} \right)^3 e^{i\mathbf{k} \cdot \mathbf{r}} k^2 \sin\vartheta_k \, dk \, d\vartheta_k \, d\varphi_k \\ &= \int_0^\infty \sum_{l=0}^\infty \sum_{m=-l}^l S_{k,l,m} \Psi_{k,l,m}(r, \vartheta, \varphi) k^2 dk, \end{aligned}$$

where $(k, \vartheta_k, \varphi_k)$ are the spherical coordinates of the wave vector $\mathbf{k}$. A plane wave in 3D can be expanded in spherical waves [5]. As a consequence, there also exists a relation of the coefficients:

$$S_{k,l,m} = \frac{i^l}{k} \int_0^\pi \int_0^{2\pi} C_{k,\vartheta_k,\varphi_k} Y_{lm}^*(\vartheta_k, \varphi_k) \sin\vartheta_k \, d\vartheta_k \, d\varphi_k.$$

Except for a constant factor, $S_{k,l,m}$ is the SH coefficient of $C_{k,\vartheta_k,\varphi_k}$ with (ϑ_k, φ_k) as the variables.

A function $f(r, \vartheta, \varphi)$ that is defined on a solid sphere of finite radius a can be expanded either in a normal Fourier series or in a spherical Fourier series. Here, the normal Fourier series is defined on a rectangular box containing the solid sphere and is centered at the origin. Let the volume of the rectangular box be V.
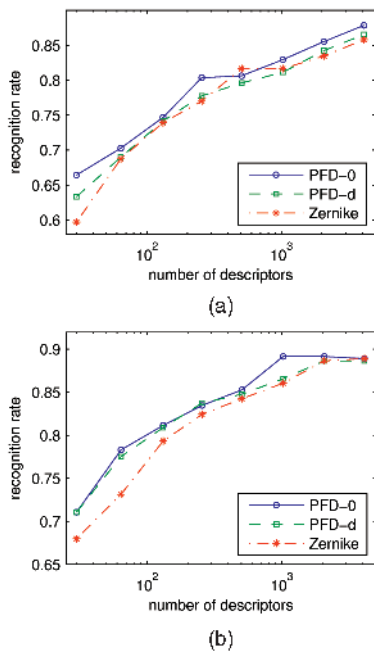
Fig. 7. Recognition rates on 2D data with (a) nearest-neighbor classifiers based on L1-norm distance and (b) support vector machines.

$$f(r, \vartheta, \varphi) = \sum_{\mathbf{k}_0} C_{\mathbf{k}_0} \frac{1}{\sqrt{V}} e^{i\mathbf{k}\cdot\mathbf{r}}$$
$$= \sum_{nlm} S_{nlm} \Psi_{nlm}(r, \vartheta, \varphi).$$

The coefficients have the following relationship if k_{nm} are selected with the zero boundary condition:

$$S_{nlm} = \left[(-1)^n i^l 4\pi \sqrt{\frac{2a}{V}} \right] \sum_{\mathbf{k}_0} \frac{k_{nl} j_l(k_0 a)}{k_0^2 - k_{nm}^2} Y_{lm}^*(\vartheta_{k_0}, \varphi_{k_0}) C_{\mathbf{k}_0}.$$

4 EXPERIMENTS

In this paper, we conduct experiments on the application of PFD and SFD to pattern classification problems. The aim here is to compare their performance with other features like Zernike descriptors or rwSHD. We shall calculate descriptors directly from the gray-value images instead of trying to get a higher classification rate by combination with other techniques that are irrelevant to the comparison, such as, texture decomposition or image warping according to the object shape.

The data set for the experiments is 3D confocal laser scanning microscopic images of pollen grains. It contains 387 images of single pollen grains from 26 species, with 13-15 objects for each species. The class of every pollen grain is known. The size of each 3D image is $128 \times 128 \times 128$, with the same resolution in each direction. Essentially, they are the same as used in [20]. We have only applied new segmentation and downscaling methods on the raw data here. As pollen grains usually have near-spherical shapes and show certain rotational symmetry, they make up a good data set for applying polar and spherical Fourier analysis. We use this data set for both 3D and 2D experiments, with the xy slice containing the object center as the 2D data.

Each image is first mapped to polar or spherical coordinates. The origin is chosen to be the center of the object, found by segmentation. For 2D, the region $\sqrt{x^2 + y^2} \leq 64$ is mapped into a rectangular grid of size 90×512 in polar coordinates (r, φ). For 3D, the region $\sqrt{x^2 + y^2 + z^2} \leq 64$ is mapped into a cuboid of size
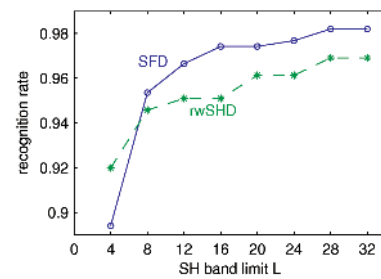

Fig. 8. Recognition rates on 3D data with nearest neighbor classifier based on L1-norm distance.

$111 \times 512 \times 512$ in spherical coordinates (r, ϑ, φ). The sizes for ϑ and φ are chosen to be the same due to the requirement of *S2Kit*, a software package for fast SH transform [26].

Due to the existence of fast programs for Fourier transform [25] and SH transform [26], the angular-part transformation can be done first. The radial part needs only to be performed for the kept terms. We use the *Gnu Scientific Library* [27] for Bessel functions and keep a lookup table for the radial functions.

4.1 2D Experiments

We test PFDs with the zero and the derivative boundary conditions. They will be called PFD-0 (0 for zero) and PFD-d (d for *d*erivative) respectively. For comparison, Zernike descriptors are also calculated, which are the magnitude of the coefficients in normalized Zernike functions. The truncation of the polar Fourier expansion is according to the energy level and the truncation of the Zernike expansion is according to the order of the polynomials. We test on descriptor sets of different sizes. Each set is normalized to have a sum of 1. As a note for the interested readers, we have $m_{max} = 170$ and $n_{max} = 57$ for 4,096 PFD-0 features.

We first perform leave-one-out tests using nearest neighbor classifiers based on the L1-norm distance. As Fig. 7a shows, PFD-0 performs best for most cases, but the difference of performance is not very distinct. We further employ support vector machines [24] for classification. The 10-fold cross-validation results are given in Fig. 7b. Although the Zernike descriptors catch up at large numbers of descriptors, they are worse when the number is low. The recognition rate of PFD-0 reaches its highest with 1,024 descriptors and is "saturated" after that. The other two sets of descriptors reach a slightly lower "saturation" level with about 2,070 or more descriptors. PFD-0, therefore, describes this data set more efficiently for the considered classification problem.

4.2 3D Experiments

For 3D, we test SFDs with the zero boundary condition against rwSHDs. These two kinds of descriptors correspond to the same angular transformation and differ in how radial patterns are represented. As the recognition rates of rwSHDs can vary with the number of sampling points n_r in radius, we compare SFDs and rwSHDs for the same SH band limits L (l is limited to $l < L$). The aim is to find whether the radial spherical Bessel transform has advantages over direct sampling.

We select the best results of each type for comparison. rwSHDs are calculated with $n_r = 8, 16, 32, 64, 111$. The numbers of SFDs are limited to powers of two between 256 and 2,048. The best results of each type for different band limits are shown in Fig. 8. Except for very low band limits, SFDs perform clearly better.

5   CONCLUSION AND OUTLOOK

We propose to use the eigenfunctions of the Laplacian that are separable in polar and spherical coordinates as basis functions for Image analysis. This idea puts the proposed polar and spherical Fourier transform and the normal Fourier transform into the same framework and ensures close resemblance and relation between them.

Rotation-invariant descriptors based on polar and spherical Fourier coefficients have been tested in object classification tasks. For 2D data, it is shown that PFDs generally perform better than Zernike descriptors for the selected data set, leaving PFDs also as candidates for applications where Zernike descriptors are used. For 3D data, SFDs perform better than rwSHDs with the same SH band limits, showing that the radial structure is effectively coded into the spherical Fourier coefficients. Also, it should be emphasized that the radial transform is automatically a multiscale approach that decomposes the pattern into different scales, just like the normal Fourier transform does.

In the experiments, the features are calculated for the whole objects. They can also be calculated pointwise to detect local structures. In that case, no mapping to the corresponding coordinates is necessary. The coefficients can be calculated by convolution, which can be done fast with the help of FFT [22]. Although the experiments are only performed on problems where rotational invariance is important, the polar and spherical Fourier analysis can also be applied to other problems like registration, where the orientation of the patterns is the essential information. For this kind of application, phases of the coefficients shall be kept as they carry the information of the angular positions.

In this paper, we are mainly concerned with the numerical precision of the calculations. The radial transform is calculated for every coefficient independently, which is surely not an efficient way. Whether fast algorithms exist for radial transforms is a question to be answered.

ACKNOWLEDGMENTS

REFERENCES

[1]   N.N. Lebedev, *Special Functions and Their Applications* (translated from Russian by R.A. Silverman), chapters 5, 6, 8. Dover, 1972.
[2]   W. Kaplan, *Advanced Calculus,* fourth ed., pp. 696-698. Addison-Wesley, 1991.
[3]   M. Born and E. Wolf, *Principles of Optics: Electromagnetic Theory of Propagation, Interference, and Diffraction of Light,* seventh ed., pp. 523-525. Cambridge, 1989.
[4]   E.W. Weisstein, "Jacobi-Anger Expansion," From MathWorld—A Wolfram Web Resource, http://mathworld.wolfram.com/Jacobi-AngerExpansion. html, 2009.
[5]   K.E. Schmidt, "The Expansion of a Plane Wave," http://fermi.la.asu.edu/ PHY577/notes/plane.pdf, 2009.
[6]   S. Erturk and T.J. Dennis, "3D Model Representation Using Spherical Harmonics," *Electronics Letters,* vol. 33, no. 11, pp. 951-952, 1997.
[7]   M. Kazhdan, T. Funkhouser, and S. Rusinkiewicz, "Rotation Invariant Spherical Harmonic Representation of 3D Shape Descriptors," *Proc. Symp. Geometry Processing,* pp. 167-175, 2003.
[8]   H. Huang, L. Shen, R. Zhang, F. Makedon, A. Saykin, and J. Pearlman, "A Novel Surface Registration Algorithm with Medical Modeling Applications," *IEEE Trans. Information Technology in Biomedicine,* vol. 11, no. 4, pp. 474-482, July 2007.
[9]   C.-H. Teh and R.T. Chin, "On Image Analysis by the Methods of Moments," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 10, no. 4, pp. 496-513, July 1988.
[10] Y. Zana and R.M. Cesar Jr., "Face Recognition Based on Polar Frequency Features," *ACM Trans. Applied Perception,* vol. 3, no. 1, pp. 62-82, 2006.
[11] A. Khotanzad and Y.H. Hong, "Invariant Image Recognition by Zernike Moments," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 12, no. 5, pp. 489-497, May 1990.
[12] W.-Y. Kim and Y.-S. Kim, "Robust Rotation Angle Estimator," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 21, no. 8, pp. 768-773, Aug. 1999.
[13] E.M. Arvacheh and H.R. Tizhoosh, "Pattern Analysis Using Zernike Moments," *Proc. IEEE Instrumentation and Measurement Technology Conf.,* vol. 2, pp. 1574-1578, 2005.
[14] A. Averbuch, R.R. Coifman, D.L. Donoho, M. Elad, and M. Israeli, "Accurate and Fast Discrete Polar Fourier Transform," *Proc. Conf. Record 37th Asilomar Conf. Signals, Systems, and Computers,* vol. 2, pp. 1933-1937, 2003.
[15] A. Makadia, L. Sorgi, and K. Daniilidis, "Rotation Estimation from Spherical Images," *Proc. 17th Int'l Conf. Pattern Recognition,* vol. 3, pp. 590-593, 2004.
[16] R. Skomski, J.P. Liu, and D.J. Sellmyer, "Quasicoherent Nucleation Mode in Two-Phase Nanomagnets," *Physical Rev. B,* vol. 60, pp. 7359-7365, 1999.
[17] B. Pons, "Ability of Monocentric Close-Coupling Expansions to Describe Ionization in Atomic Collisions," *Physical Rev. A,* vol. 63, p. 012704, 2000.
[18] R. Bisseling and R. Kosloff, "The Fast Hankel Transform as a Tool in the Solution of the Time Dependent Schrödinger Equation," *J. Computational Physics,* vol. 59, no. 1, pp. 136-151, 1985.
[19] D. Lemoine, "The Discrete Bessel Transform Algorithm," *J. Chemical Physics,* vol. 101, no. 5, pp. 3936-3944, 1994.
[20] O. Ronneberger, E. Schultz, and H. Burkhardt, "Automated Pollen Recognition Using 3D Volume Images from Fluorescence Microscopy," *Aerobiologia,* vol. 18, pp. 107-115, 2002.
[21] H. Burkhardt, *Transformationen zur lageinvarianten Merkmalgewinnung,* Ersch. als Fortschrittbericht (Reihe 10, Nr. 7) der VDI-Zeitschriften. VDI-Verlag, 1979.
[22] O. Ronneberger, J. Fehr, and H. Burkhardt, "Voxel-Wise Gray Scale Invariants for Simultaneous Segmentation and Classification," *Proc. 27th DAGM Symp.,* pp. 85-92, 2005.
[23] Q. Wang, O. Ronneberger, and H. Burkhardt, "Fourier Analysis in Polar and Spherical Coordinates," internal report, http://lmb.informatik.unifreiburg.de/papers/, 2008.
[24] C.C. Chang and C.J. Lin, "LIBSVM—A Library for Support Vector Machines," http://www.csie.ntu.edu.tw/cjlin/libsvm/, 2009.
[25] FFTW Home Page, http://www.fftw.org/, 2009.
[26] Fast Spherical Harmonic Transforms, http://www.cs.dartmouth.edu/ geelong/sphere/, 2009.
[27] The Gnu Scientific Library, http://www.gnu.org/software/gsl/, 2009.

▷ **For more information on this or any other computing topic, please visit our Digital Library at** www.computer.org/publications/dlib.