Computer Vision and Image Understanding
Vol. 74, No. 1, April, pp. 22–35, 1999
Article ID cviu.1999.0745, available online at http://www.idealibrary.com on IDEAL®

Pattern Matching as a Correlation on the Discrete Motion Group

Alexander B. Kyatkin and Gregory S. Chirikjian*

Department of Mechanical Engineering, Johns Hopkins University, 3400 North Charles Street, Baltimore, Maryland, 21218
E-mail: gregc@jhu.edu

Received November 3, 1997; accepted January 6, 1999

In this paper we develop a correlation method for the template matching problem in pattern recognition which includes translations, rotations, and dilations in a natural way. The correlation method is implemented using Fourier analysis on the "discrete motion group" and fast Fourier transform methods. A brief introduction to Fourier methods on the discrete motion group is given and the efficiency of these methods is discussed. Results of the numerical implementation are given for particular examples.

Key Words: **pattern analysis; object recognition and indexing.**

1. INTRODUCTION

In this paper we address a two-dimensional problem in pattern recognition. For a given template object we want to find if this template object is present in a given image, and, if it is found, determine its position and orientation. We use a correlation method (see [1] and references therein) for this purpose, which is extended in a natural way to include rotations and dilations of the template object in addition to translations. Essentially, we translate, rotate, and dilate the template object, overlap it with the image and compute an overlap area (weighted by the intensity value at each pixel) with the proper normalization. The novelty of our approach is that the correlation method is implemented using the Fourier transform on the "discrete motion group." Fourier methods on the discrete motion group also provide a fast method to distinguish "identical" images (up to possible translations and rotations of the image) from "different" ones.

The discrete motion group can be viewed as the set of matrices of the form

$$g = \begin{pmatrix} R & \mathbf{r} \\ \mathbf{0}^T & 1 \end{pmatrix}, \tag{1}$$

where

$$R = \begin{pmatrix} \cos 2\pi i/N & -\sin 2\pi i/N \\ \sin 2\pi i/N & \cos 2\pi i/N \end{pmatrix} \tag{2}$$

for fixed natural number N and $i \in [0, N-1]$ ($R \in SO(2)$ for the continuous motion group $SE(2)$).[1] The group law is simply matrix multiplication.

The problem of template matching is quite old and has been approached in a number of different ways. Perhaps the most common (and oldest) approach is that of "matched filters" [2]. In this approach the Fourier transform of the image and template are taken, these are multiplied, and a peak is sought. This method can be implemented via digital computer, or by analog optical computation [3]. The drawback of this standard approach is that rotations are handled in a very awkward manner. Several works have considered rotation-invariant approaches (e.g., [4]). In such approaches, polar coordinates are used and images are expanded in series of Zernike polynomials (see, e.g. [5]) or by using the Hankel transform. The problem with such approaches is that rotational invariance is often gained at the expense of the translational invariance offered by the classical Fourier transform.

A number of works have considered using invariants of images for recognition (e.g. [6]). When one begins discussing invariants, the most natural analytical tool is group theory. In this work we apply an area of group theory called noncommutative harmonic analysis to the template matching problem. In short, this area of mathematics deals with the generalization of the concept of convolution and Fourier transforms to functions on groups. In particular, if we are given a function $f(\mathbf{x})$, the generalized Fourier transform developed and applied in this paper is a matrix function which has the property

$$\mathcal{F}(f(R^T(\mathbf{x}-\mathbf{a}))) = \mathcal{F}(f(\mathbf{x}))U(R, a),$$

where U is a unitary matrix that depends on rotation R and translation $\mathbf{a}$, and $\mathcal{F}$ denotes the nonabelian Fourier transform. The above expression cannot be written as a matrix product for the usual abelian Fourier transform for $R \neq$ identity, although it is completely analogous to the behavior of the abelian Fourier transform applied to translated functions. In other words, noncommutative harmonic analysis provides a natural tool for

* To whom all correspondence should be addressed.

[1] The notation $SE(2)$ stands for "special Euclidean" group of $\mathbb{R}^2$, i.e. the group of all rigid-body motions in the plane. It is also called the Euclidean motion group, or simply the motion group.

1077-3142/99 $30.00

translation *and* rotation invariant pattern matching. Furthermore, since U is unitary $\|U\mathcal{F}(f)\|_2 = \|\mathcal{F}(f)\|_2$, and so this generalized Fourier transform provides a tool for generating a whole continuum of pattern invariants under rigid-body motion.

The connection between group theory and the theory of wavelets (which has become a very popular tool in image analysis) has been well established. In essence, expanding a function in a wavelet basis is achieved by starting with a mother wavelet and superposing affine-transformed versions of the mother wavelet to best approximate a given function. The interested reader is pointed to [7–10] for further reading on the subject of wavelets, their applications in image analysis, and their connection with group theory.

The approach presented in this paper is to use the nonabelian Fourier transform and generalized concepts of convolution and correlation. This is very different than wavelet approaches. While wavelets typically allow one to efficiently approximate functions (or images), they have the drawback of not behaving well under operations such as convolution, which is the most natural tool in matched filtering.

In Section 2 we describe the correlation method. Section 3 is an introduction to Fourier analysis on the discrete motion group. In Section 4 we describe the implementation of the correlation method using the Fourier transform on the discrete motion group. Section 5 describes the invariant constructions which may be used in image analysis problems. Section 6 examines the computational complexity of the approach. Section 7 describes practical numerical examples: Subsection 7.1 gives numerical examples of the correlation method which includes translations and rotations; Subsection 7.2 illustrates applications of the invariants on the discrete motion group for comparison of images.

2. METHOD FOR PATTERN RECOGNITION

In this paper we extend the correlation method for pattern recognition [1] to include, in a natural way, rotations and dilations (in addition to translations) as the allowed transformations of the image. To find if the template object is present in the image we take a section from the image and compare it with a rotated, translated, and dilated version of the template pattern. Taking a section from the image is equivalent to multiplication of the image by a "window" function, which is rotated, translated, and dilated the same way as the template pattern.

Mathematically the correlation function is written as

$$q(\mathbf{a}, R, k) = \frac{\int_{\mathbb{R}^2} f_1(\mathbf{x}) W(R^{-1}(k\mathbf{x}-\mathbf{a}))\, f_2(R^{-1}(k\mathbf{x}-\mathbf{a}))\, d^2x}{\left[\int_{\mathbb{R}^2} (f_1(\mathbf{x}))^2 (W(R^{-1}(k\mathbf{x}-\mathbf{a})))^2\, d^2x\right]^{1/2} \int_{\mathbb{R}^2} (f_2((R^{-1}(k\mathbf{x}-\mathbf{a})))^2\, d^2x]^{1/2}}, \tag{3}$$

where $R \in SO(2)$, $\mathbf{a} \in \mathbb{R}^2$, $k \in \mathbb{R}^+$ close to one, and $W(\mathbf{x})$ is a window function. For a similar template pattern and window from the image the value of the correlation coefficient should be close to one. We note that the integral

$$\int_{\mathbb{R}^2} (f_2(R^{-1}(k\mathbf{x}-\mathbf{a})))^2\, d^2x \tag{4}$$

is just the square of norm of function f_2 (for $k = 1$),

$$\int_{\mathbb{R}^2} (f_2(\mathbf{x}))^2\, d^2x.$$

According to the Cauchy–Schwarz inequality,

$$\int f_1(\mathbf{x})\, f_2(\mathbf{x})\, d^2x \leq \left[\int (f_1(\mathbf{x}))^2\, d^2x \int (f_2(\mathbf{x}))^2\, d^2x\right]^{1/2},$$

the correlation coefficient (3) is always smaller or equal to one, and it is equal to one for an identical pattern and windowed image. We note that the value of a correlation coefficient does not change if we change overall intensity of the original image or template object.

For a dilation coefficient $k = 1$ we observe that the correlation function $q = q(\mathbf{a}, R)$ is a function on the Euclidean motion group $SE(2)$ [22, 16], which is the semidirect product of translation group $(\mathbb{R}^2, +)$ and the rotation group $SO(2)$. It appears that this group has not been used extensively in applications to the image processing; the authors are aware of only a few previous works using this group (e.g., [11, 12, 15]).

Using Fourier methods on the motion group we compute the correlation coefficient in a much more efficient way than using direct integration. Indeed, the direct computation of integral (3) is very costly (we consider for simplicity the $k = 1$ case). For $N_r = N_x \cdot N_y$ samples of the image (and template) on an $N_x \times N_y$ rectangular grid, and for N samples of orientation, we need to perform $O(N_r^2 N)$ computations (and we need to compute the convolution-like integrals twice, in the denominator and numerator of (3)). For $N_r = 256 \times 256$ and $N = 60$, the computations require 5×10^{11} operations, which requires a day of computer work on a 250 MHz workstation. In this paper we use the advantages of Fourier methods on the "discrete motion group" (i.e. the subgroup of $SE(2)$, where the orientation angle has discrete values from the C_N subgroup of $SO(2)$, $\theta = 2\pi i/N$ for $i = 0, \ldots, N-1$), and fast Fourier transform (FFT) methods [17, 18] to compute the correlation coefficient in $O(N\ N_r\ \log N_r)$ computation. In addition, Fourier methods on the discrete motion group provide a very fast method for comparison of two images which are translated and rotated relatively to each other.

A natural question to ask is how the computational requirements of this approach compare to classical Fourier techniques applied to matched filtering. The answer is that they are on the same order. However, the benefit of our formulation is that it provides a clean notation in which to treat translations and rotations in a unified way. This paper also serves as an introduction of the image understanding community to techniques which are not widely known outside of pure mathematics. In the next section we discuss briefly Fourier methods on the discrete motion group.

3. FOURIER TRANSFORM ON THE DISCRETE MOTION GROUP

The concept of convolution of functions on a wide variety of abstract groups is well known in the pure mathematics literature [14]. A detailed study of the concrete case of convolution of functions on $SE(2)$ in the context of robot kinematics can be found in [13].

Building on this previous work, we note that the numerator (and denominator) in the correlation function (3) for $k=1$ may be written formally as a convolution-like integral on the Euclidean motion group

$$\begin{aligned}&\int_{\mathbb{R}^2} f_1(\mathbf{x})\, f_2(R^{-1}(\mathbf{x}-\mathbf{a}))\, d^2x\\ &\quad= \int_{SO(2)}\int_{\mathbb{R}^2} \tilde{f}_1(\mathbf{x}, A)\, \tilde{f}_2(R^{-1}(\mathbf{x}-\mathbf{a}), R^{-1}\circ A)\, d^2x\, dA\\ &\quad= \int_{SE(2)} \overline{\tilde{f}_1(h)}\, \tilde{f}_2(g^{-1}\circ h)\, d\mu(h),\end{aligned} \tag{5}$$

where $A \in SO(2)$ and $dA = d\theta/(2\pi)$ is the normalized integration measure on $SO(2)$, and in the template-matching problem functions $\tilde{f}_{1,2}$ are explicit functions only of position, i.e., $\tilde{f}_{1,2}(\mathbf{x}, A) = f_{1,2}(\mathbf{x})$ (henceforth we do not distinguish between $\tilde{f}_{1,2}$ and $f_{1,2}$).[2] Furthermore, $f_{1,2}(\mathbf{x})$ are nonnegative real functions (we formally write f_1 as the complex conjugate of itself to use the properties of the Fourier transform later).

The group elements g, h are in $SE(2)$, the group product is the group product on the motion group,[3] and $d\mu(h) = d^2x\, d\theta/(2\pi)$.

We assume that the orientation angles are restricted to values from the discrete subgroup C_N of the rotation group $SO(2)$. We refer to the subgroup of the motion group with a discrete range of allowed rotations as the discrete motion group G_N.

For the discrete motion group the integration over orientation should be replaced by summation through A_i (which can be viewed as elements of the group C_N, or as matrices of the form in Eq. (2)):

$$\int_{SO(2)} (\cdot)\, dA \to \frac{1}{N}\sum_{i=0}^{N-1} (\cdot).$$

In the case of the translation group the usual Fourier transform on $\mathbb{R}^2$ may be used to get a simple expression for convolution in Fourier space (i.e. the product of Fourier transforms). In fact, this property is based on the property

$$\mathcal{U}(\mathbf{a};\mathbf{p})\cdot\mathcal{U}(\mathbf{b};\mathbf{p}) = \mathcal{U}(\mathbf{a}+\mathbf{b};\mathbf{p})$$

[2] Any function on $\mathbb{R}^2$ can also be viewed as one on $SE(2)$ which is constant over all orientations.

[3] For $g = (\mathbf{x}, R)$ and $h = (\mathbf{y}, A)$ the group product is defined as $g \circ h = (R\mathbf{y} + \mathbf{x}, R \circ A)$, where $R \circ A$ is a group product for $SO(2)$.

of the Fourier transform elements $\mathcal{U}(\mathbf{a};\mathbf{p}) = \exp(i\mathbf{p}\cdot\mathbf{a})$, which form a complete and orthonormal set of elements for all possible values of the Fourier parameter vector $\mathbf{p}$. We note that $\mathcal{U}(\mathbf{a};\mathbf{p})$ are matrix elements (complex numbers in this case) of unitary irreducible representations [25, 26] of the translation group of $\mathbb{R}^2$.

We use a similar approach in order to get a simple expression for the convolution integral on the motion group in Fourier space. We have to use a generalized Fourier transform with the property that (see Appendix)

$$\mathcal{F}(f_1 * f_2) = \mathcal{F}(f_1)\mathcal{F}(f_2),$$

where $f_{1,2}$ are functions on the motion group. This will provide a tool for fast calculation of integrals like those in Eq. (5).

A well-developed theory for such generalizations of the Fourier transform exists. It is called noncommutative harmonic analysis. A key element of this theory is the enumeration of linear operators, U, which have the homomorphism property

$$U(g_1;\rho)U(g_2;\rho) = U(g_1\circ g_2;\rho), \tag{6}$$

where $g_{1,2}$ are group elements of a group G, ρ is a generalized Fourier parameter (or set of parameters), and the operator product may be understood as a matrix product (of, in general, infinite dimensional matrices).

This homomorphism property allows one to reduce the convolution integrals to a matrix product equation in Fourier space. The property (6) is just part of the definition of a group representation [25] and is required to define Fourier transforms with the convolution property. The operators U can be thought of as generalizations of the complex exponentials, $\mathcal{U}$, used in usual Fourier analysis. Each U can be expressed as a unitary matrix.

To generate the complete and orthonormal basis in which to expand functions on the group, we have to calculate the matrix elements of *irreducible* and *unitary* representations (IURs) [25, 26] of the group. A detailed review of the general theory is provided is the Appendix.

The elements of the U matrices for the discrete motion group may be written as (see Appendix for details)

$$U_{mn}(g;\rho) = U_{mn}(A_j, \mathbf{r}; p, \phi) = e^{ip\mathbf{u}_m^{\phi}\cdot\mathbf{r}}\delta_{A_j^{-1}\mathbf{u}_m,\mathbf{u}_n}, \tag{7}$$

where A_j^{-1} is the inverse of the discrete rotation A_j, and $\mathbf{u}_k^{\phi}$ denotes the vector to the angle $\theta = \phi + 2\pi k/N$ on the unit circle in the interval $F_k = [2\pi k/N, 2\pi(k+1)/N]$, $k = 0, \ldots, N-1$ (ϕ measures the angle on this segment, $0 \le \phi \le 2\pi/N$). The vector $\mathbf{u}_k^{\phi}$ is analogous to the 2D Fourier vector $\mathbf{p}$ in ordinary 2D Fourier transform (normalized to unit magnitude) and, thus, has a dependence on the continuous angle θ, which just measures the polar angle of $\mathbf{p}$. We note that each element of the discrete motion group can be expressed as a pair $g = (A_i, \mathbf{r})$ and each Fourier parameter can be expressed as the pair $\rho = (p, \phi)$.

The direct Fourier transform is defined as

$$\hat{f}_{mn}(p,\phi) = [\mathcal{F}(f)]_{mn} = \sum_{i=0}^{N-1} \int_{\mathbb{R}^2} f(A_i, \mathbf{r}) U^{-1}_{mn}(A_i, \mathbf{r}; p, \phi)\, d^2 r. \quad (8)$$

The vector $\mathbf{u}^{\phi}_m$, which is inside the segment F_m may be received by the rotation A_m (which transform F_0 to F_m) from $\mathbf{u}^{\phi}_0$, $\mathbf{u}^{\phi}_m = A_m \mathbf{u}^{\phi}_0$. The parameter ϕ denotes the position inside the segment F_0.

The inverse Fourier transform is

$$\mathcal{F}^{-1}(\hat{f}) = \frac{1}{4\pi^2} \sum_m \sum_n \int_0^{\infty} \int_0^{2\pi/N} \hat{f}_{mn}(p,\phi) U_{nm}(A_i, \mathbf{r}; p, \phi) p\, dp\, d\phi, \quad (9)$$

where the angle ϕ is measured from $\theta = 2\pi n/N$. We note that this result is in agreement with [11].

4. APPLICATION TO THE CORRELATION METHOD

The convolution-like integrals in the numerator and denominator of (3) may be formally written (for simplicity we consider the $k = 1$ case) as integrals

$$c(\mathbf{x}, A_j) = \frac{1}{N} \sum_{i=0}^{N-1} \int_{\mathbb{R}^2} f_1(\mathbf{y}, A_i)\, f_2\big(A_j^{-1}(\mathbf{y}-\mathbf{x}),\, A_j^{-1} A_i\big)\, d^2 y \quad (10)$$

(where the functions $f_{1,2}$ are orientation-independent).

The correlation function, however, is a function on G_N, so we may use the Fourier transform on the discrete motion group G_N to write this integral as a product of Fourier transforms. Because the functions are real the integral in the numerator of Eq. (3) may be written as

$$\frac{1}{N} \sum_{i=0}^{N-1} \int_{\mathbb{R}^2} \overline{f_1(\mathbf{y}, A_i)}\, f_2\big(A_j^{-1}(\mathbf{y}-\mathbf{x}),\, A_j^{-1} A_i\big)\, d^2 y = \int_{G_N} \overline{f_1(h)}\, f_2(g^{-1}h)\, d\mu(h),$$

where we denote integration $d\mu(h)$ over the discrete motion group, G_N, to mean integration over $\mathbb{R}^2$ and summation through the A_i, and the group elements are of the form $g = (\mathbf{x}, A_j)$. Using the orthogonality and homomorphism properties of the Fourier matrix elements, this integral may be written as

$$\frac{1}{N} \sum_q \sum_n \int_0^{\infty} \int_0^{2\pi/N} \sum_m \big(\overline{\hat{f}_{1mn}}\, \hat{f}_{2mq}\big) U_{qn}(g^{-1}; p, \phi) p\, dp\, d\phi = \frac{1}{N} \sum_q \sum_n \int_0^{\infty} \int_0^{2\pi/N} \sum_m \big(\overline{\hat{f}_{2mq}}\, \hat{f}_{1mn}\big) U_{nq}(g; p, \phi) p\, dp\, d\phi, \quad (11)$$

where ϕ is measured from $2\pi q/N$. For the second expression we used the unitarity of the matrix elements U_{mn} and the fact that the expression is real (i.e., we take the complex conjugate of the integral). The matrices $(\hat{f}_{1,2})_{mn}$ are the Fourier transforms (as defined in Eq. (8)) of the functions $f_{1,2}(\mathbf{x}, A_i)$. We note that this integral is the inverse Fourier transform of $\hat{f}^{\dagger}_2 \cdot \hat{f}_1$, and thus, the expression depends only on three indices.

Because functions $f_{1,2}(\mathbf{x}, A_i) = f_{1,2}(\mathbf{x})$ do not depend on the orientations A_i, matrix elements in the same column are the same, i.e.,

$$(\hat{f}_{1,2})_{mn} = (\hat{f}_{1,2})_{qn}$$

for any m, q. This may be observed from the expression

$$U_{mn}(g^{-1}; p, \phi) = e^{-ip\mathbf{u}^{\phi}_n \cdot \mathbf{r}} \delta_{A_i^{-1}\mathbf{u}_n, \mathbf{u}_m}$$

(the exponent depends only on the n-index), the definition of the direct transform (8), and the fact that the functions do not depend on the orientation. Thus, we compute a row of the Fourier matrix for a particular orientation (for example $A_0 = \mathbf{1}$, the identity element)

$$(\hat{f}_{1,2})_n = (\hat{f}_{1,2})_{nn}.$$

This may be done using the 2D FFT for the functions $f_{1,2}(\mathbf{x})$ and interpolating the Foureir values to points on a polar coordinate grid.

The value of p is determined by $|\mathbf{p}|$, the values of m and v are determined by the angular part of $\mathbf{p}$. This requires $O(N_r \log(N_r))$ computations.

Thus, the integrals in Eq. (3) may be written as

$$c(\mathbf{x}, A_j) = C \sum_q \sum_n \int_0^{\infty} \int_0^{2\pi/N} \big(\overline{\hat{f}_{2q}(p,\phi)}\, \hat{f}_{1n}(p,\phi)\big) U_{nq}(g; p, \phi) p\, dp\, d\phi, \quad (12)$$

where $C = 1/N$.

We observe that the convolution-like integrals may be computed by taking the Fourier transform, computing the product of transforms, and taking the inverse Fourier transform on the discrete motion group.

5. INVARIANTS OF THE DISCRETE MOTION GROUP

Let us assume that one wants to compute properties of the image (object) which are invariant with respect to translations and rotations of the image. The Fourier transform on the discrete motion group provides a very efficient tool to compute these invariants. Let us construct a function with values in $\mathbb{R}^+$,

$$\eta(p;\phi) = \sum_{m=0}^{N-1} [\overline{\hat{f}_m(p;\phi)}\, \hat{f}_m(p;\phi)], \quad (13)$$

for each fixed $\phi = 0, \ldots, N_\phi - 1$, where $\hat{f}_m(p;\phi)$ is the Fourier transform on the discrete motion group of $f(\mathbf{x})$. Then (13) is invariant with respect to rotations and translations of $f(\mathbf{x})$; i.e., $\eta(p;\phi)$ does not change if we compute (13) using the Fourier transform on the motion group for $f'(\mathbf{x}) = f(R^{-1}(\mathbf{x} - \mathbf{a}))$.

We note that for orientation-independent functions (i.e., for functions on $\mathbb{R}^2$) the Fourier transform elements $\hat{f}_m$ may be arranged as a matrix which has the same matrix elements in the same column,

$$\hat{f}_{qm} = \hat{f}_{rm} = \hat{f}_m.$$

Then (13) may be written also as a trace

$$\eta(p;\phi) = \mathrm{Tr}[\hat{f}^\dagger(p;\phi)\hat{f}(p;\phi)], \tag{14}$$

where $\hat{f}^\dagger(p;\phi)$ is the Hermitian conjugate matrix.

According to (8), $\hat{f}_{qm}$ for $f(\mathbf{x})$ may be written as

$$\hat{f}_{qm}(p;\phi) = \int_{G_N} f(h) U^{-1}_{qm}(h; p, \phi)\, d\mu(h),$$

where the integral over G_N denotes integration with respect to $\mathbf{x}$ *and* summation through the elements of C_N, and $f(h) = f(\mathbf{x})$. The function $f'(\mathbf{x}) = f(R^{-1}(\mathbf{x} - \mathbf{a}))$ may be formally written as $f(g^{-1} \circ h)$, where $g = (\mathbf{a}, R) \in G_N$. Then $\hat{f}'_{qm}$ is written as

$$\hat{f}'_{qm}(p;\phi) = \int_{G_N} f(g^{-1} \circ h) U^{-1}_{qm}(h; p, \phi)\, d\mu(h).$$

Using the invariance of the integration measure we write this integral as

$$\hat{f}'_{qm}(p;\phi) = \int_{G_N} f(h') U^{-1}_{qm}(g \circ h'; p, \phi)\, d\mu(h').$$

Using the homomorphism properties of U we write it as

$$\begin{aligned}\hat{f}'_{qm}(p;\phi) &= \left[\int_{G_N} f(h') U^{-1}_{qr}(h'; p, \phi)\, d\mu(h')\right] \cdot U^{-1}_{rm}(g; p, \phi)\\ &= \hat{f}_{qr}(p, \phi) U^\dagger_{rm}(g; p, \phi),\end{aligned}$$

where we have used a unitarity property of U. Thus, the Fourier matrix is transformed under rotations and translations $g \in G_N$ as

$$\hat{f}'(p, \phi) = \hat{f}(p, \phi) U^\dagger(g; p, \phi).$$

Using the cyclic property of Tr and unitarity of U it is clear that

$$\begin{aligned}&\mathrm{Tr}[(\hat{f}')^\dagger(p, \phi)\hat{f}'(p, \phi)]\\ &\quad= \mathrm{Tr}[U(g; p, \phi)\hat{f}^\dagger(p, \phi)\hat{f}(p, \phi)U^\dagger(g; p, \phi)]\\ &\quad= \mathrm{Tr}[\hat{f}^\dagger(p, \phi)\hat{f}(p, \phi)],\end{aligned}$$

which proves the invariance of (13). We note that the invariant, written in the form (14), is valid also for orientation-dependent functions (i.e. for general functions on the discrete motion group). The use of invariants for pattern recognition was suggested in [11].

6. EFFICIENT CALCULATION OF CONVOLUTION-LIKE INTEGRALS USING THE FOURIER METHOD

As we mentioned before, the direct integrations of (3) requires $O(N_r^2 N)$ computations for C_N, where N_r is the number of sampling points in an $\mathbb{R}^2$ region.

Using the Fourier transform on the discrete motion group we have to compute direct Fourier transforms for image and template, compute the matrix product (in our case it is a column–row product) of the Fourier transform, which describes the Fourier transform of the convolved functions, and then calculate the inverse Fourier transform.

The calculation of direct Fourier transform and the "matrix" (column–row) product is a fast computation. The direct Fourier transform for $f_{1,2}(\mathbf{x})$ may be computed using a usual two-dimensional FFT [18] in $O(N_r \log N_r)$ computations. The FFT gives, however, values of Fourier elements computed on the Cartesian square (rectangular) grid of $\mathbf{p}$ values. To receive the Fourier transform elements $\hat{f}_m(p, \phi)$ on the discrete motion group we have to interpolate values on the Cartesian grid to a polar coordinate grid, the p value is the magnitude of $\mathbf{p}$, the m and ϕ indices are determined by the angular part of $\mathbf{p}$ (thus, the constraint $N_p N_\phi N \approx N_r$ may be used). The linear interpolation requires $O(N_r)$ computations. The product of Fourier column $\overline{\hat{f}^T_m(p, \phi)}$ and row $\hat{f}_n(p, \phi)$ which gives Fourier matrix $\hat{F}_{mn}(p, \phi)$, may be performed in $O(N^2 N_p N_\phi) = O(N N_r)$ computations. We note that the trace in invariants (13) may be computed in $O(N N_p N_\phi) = O(N_r)$ (for all ϕ-values).

Thus, the direct Fourier transform and the "matrix" product may be computed in $O(N_r \log N_r + O(N N_r)$ computations.

The inverse Fourier transform calculation is a slower computation. One element from each row and column of $\hat{F}_{mn}(p, \phi)$ is used in computation of the inverse Fourier transform for each rotation element A_i. First, we interpolate the value of the Fourier transform on the square grid $N_r \times N_r$ of $\mathbf{p}$ to polar coordinates. The radial coordinate is $p = |\mathbf{p}|$, the polar angle is determined by the values of m and ϕ (the value of n is determined by m and the index of rotation i, $n = m + i$, thus we take $\hat{F}_{m,m+i}$ elements from the Fourier matrix to compute the inverse transform for fixed orientation A_i). After inverse interpolation to Cartesian coordinates (which may be done in $O(N N_r)$ computations), the inverse Fourier integration may be performed in $O(N_r \log(N_r))$ for each of the N nonzero matrix elements of U using the FFT. Thus, in $O(N N_r \log(N_r))$ computations we reproduce the function for all A_i. We note that the inverse Fourier transform computation is $O(N)$ (or $O(\log N_r)$, depending which is larger) times more time-consuming, because we reproduce a function on the discrete motion group, rather than a function on $\mathbb{R}^2$.

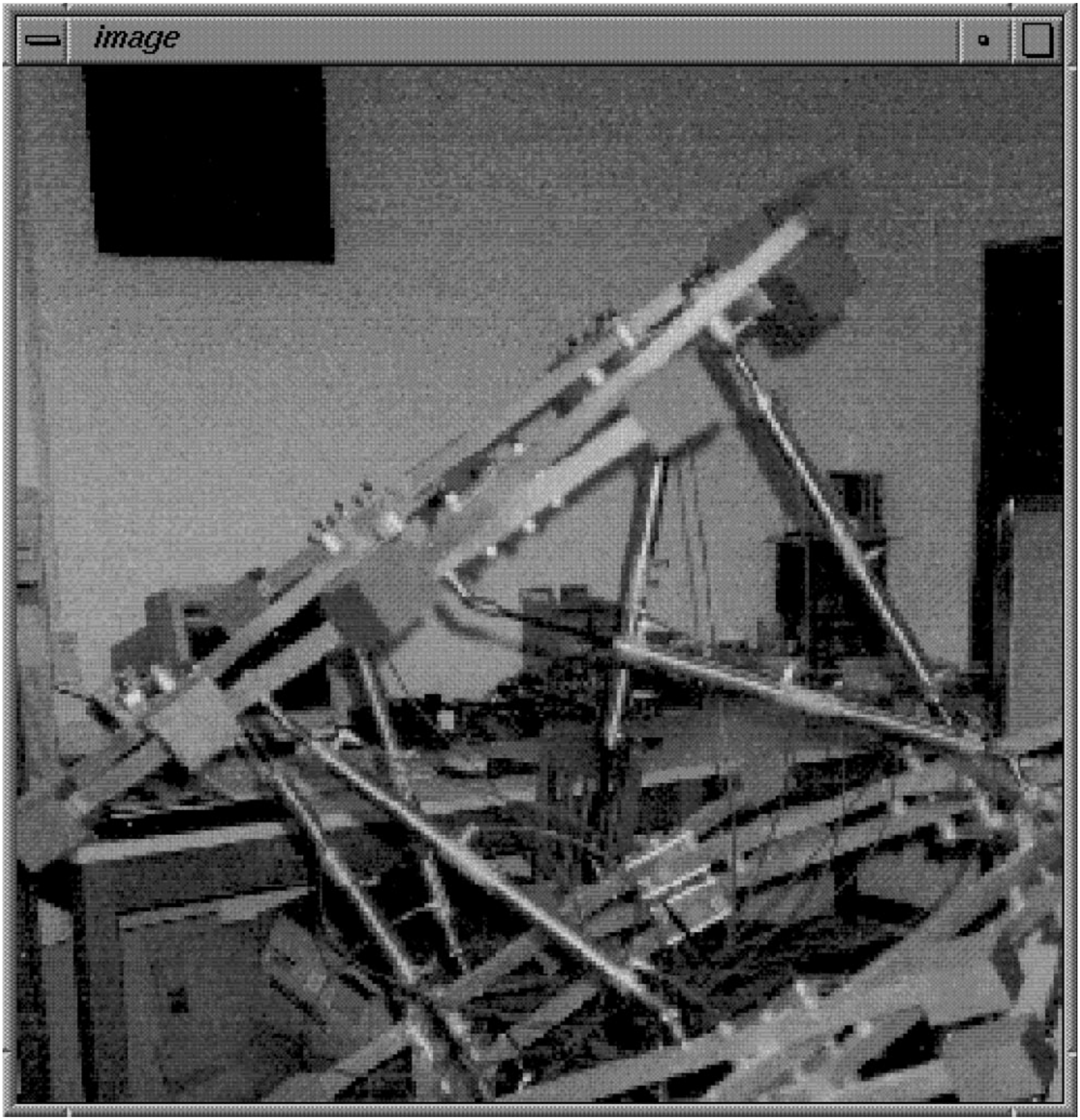

FIG. 1. The image—256 × 256 array of grey values.

Thus, the total required is $O(N N_r \log(N_r))$ computations, and these computations are, for the most part, calculations of the inverse Fourier transform. We also note that we have to perform calculations twice, to compute convolution-like integrals in the denominator and numerator of (3).

The total order of computations using classical Fourier analysis is of the same order as in our implementation, but the Fourier transform on the motion group has additional nice properties. For example, it allows one to neatly write integrals in the correlation function as matrix products in Fourier space. It also gives an efficient way to construct image invariants. These invariants may be used in image processing problems (see Section 7.2 for numerical examples). Of course it can be argued that the classical (scalar) Fourier transforms of rotated versions of images can be arranged to form a Fourier matrix like ours. If this arrangement is performed, then knowingly or not, one is calculating the Fourier transform on the discrete motion group.

FIG. 2. The template pattern. The arrow is used as a reference to find the position and orientation of this pattern in the image.

7. NUMERICAL EXAMPLES

7.1. Correlation Method, Including Rotations and Translations

In this section we compute the correlation function, Eq. (3) (for dilation $k = 1$) for some practical examples. We compute most examples for $N_r = 256 \times 256$ and $N = 60$ (C_{60} group), although the computing time for other arrays is also reported.

We consider the image depicted in Fig. 1. This is a 256×256 array of grey values (256 grey levels of intensity for each pixel). We choose a template, depicted in Fig. 2, which is a rotated (at angle $\theta = -\pi/3$), and the translated pattern taken from the image. The arrow shown on the picture is used as a reference arrow to find the position and orientation of this template in the image. The correlation function depicted for the $\theta = \pi/3$ angle is depicted in Fig. 3. The highest value of the correlation function is at the original position and orientation of the pattern in the image. We also find positions and orientations of local maxima in each of $m \times m$ subregions of the original image. For $m = 8$ the

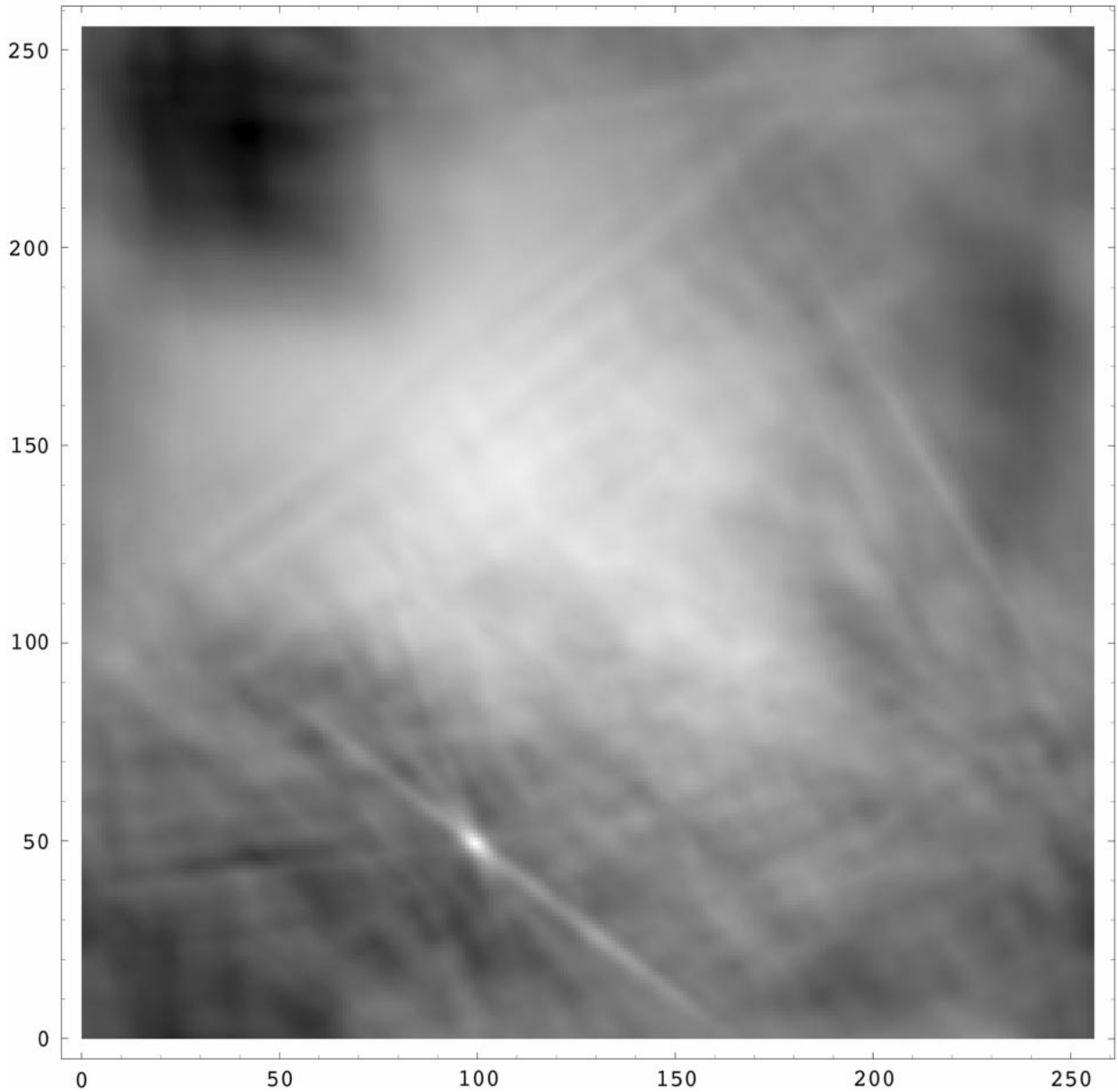

FIG. 3. The correlation function depicted for the $\theta = \pi/3$ orientation angle.

positions and orientations of local maxima in each of subregions are shown in Fig. 4. The highest value is depicted by the arrow, which is rotated and translated from the arrow in Fig. 2. Other local maxima (with a value of correlation which is greater than 0.85) are depicted by a white square; a small line attached to the square shows the orientation.

We note that the precise values of correlation at the locations of 64 maxima may be found by direct integration, and the Fourier method may be used as a fast filter method to find locations of these maxima. It is especially important to compute precise values in the case when the template object does not match exactly the pattern in the image.

In the table below we listed the computing time of the method (given in minutes and seconds, on a 250-MHz SGI workstation), implemented in the C programming language. N is listed along the horizontal, the right column lists the time to compute the correlation coefficients at 64 maxima using direct integration. The N_r array size is given along the vertical.

	$N = 60$	$N = 30$	$N = 10$	Dir. int.
$N_r = 256 \times 256$	4:06	2:11	0:48	0:55
$N_r = 128 \times 128$	1:12	0:36	0:16	0:13
$N_r = 64 \times 64$	0:25	0:12	0:04	0:03

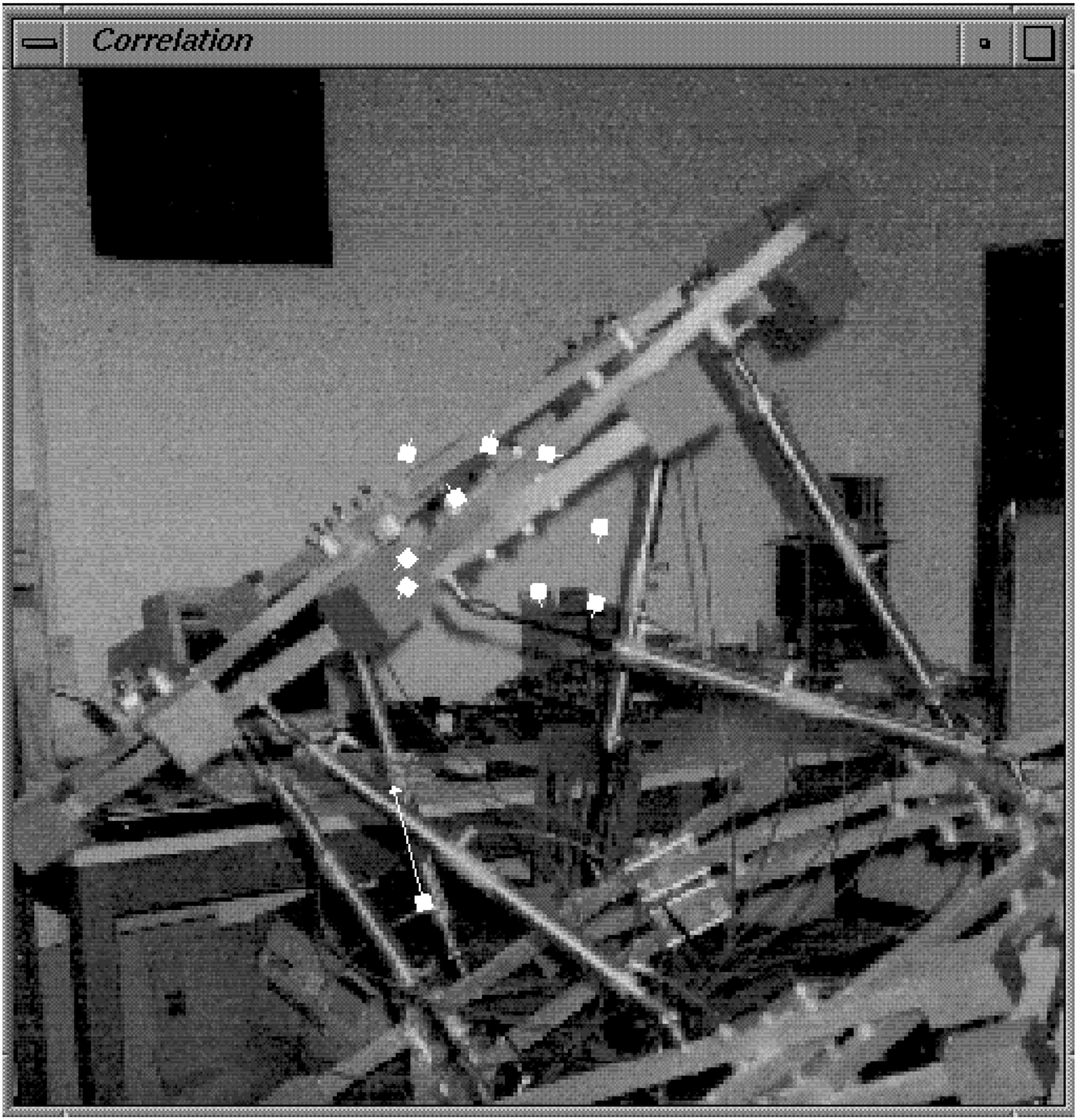

FIG. 4. The image. Positions and orientations of the absolute maximum (shown by the arrow) and local maxima of the correlation function are depicted.

7.2. Using the Invariants on the Motion Group to Compare Images

As we have shown before, function (13) is the same for images which are rotated and translated relative to each other. It may be used to compare images and determine if they are identical (similar) or not. Again, we compute the correlation coefficient of invariant functions $f_{1,2}(p,\phi)$, computed for two different images,

$$\eta(\phi) = \frac{\int_0^\infty f_1(p,\phi)\, f_2(p,\phi)\, p\, dp}{\left(\int_0^\infty f_1(p,\phi)^2\, p\, dp\right)^{1/2} \left(\int_0^\infty f_2(p,\phi)^2\, p\, dp\right)^{1/2}}.$$

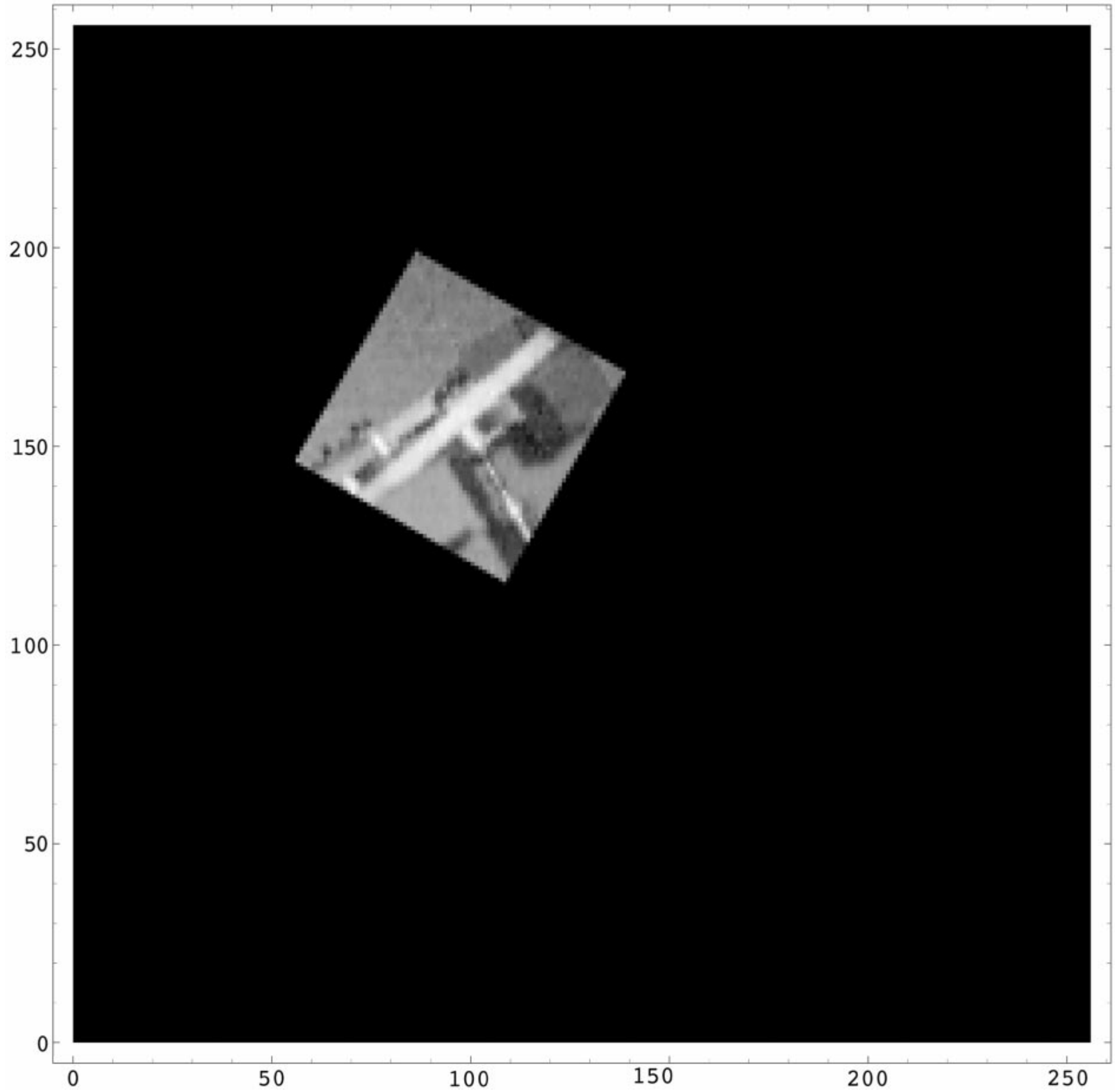

FIG. 5. An image.

This is a fast computation of the order $O(N_p N_\phi) \approx O(N_r/N)$ (to compute N_ϕ coefficients) and it may be done using the usual integration techniques. As we mentioned before, the direct Fourier transform may be computed in $O(N_r \log(N_r))$ computation; computation of the sum in (13) may be done in $O(N_r)$ computations.

We compare the images depicted in Fig. 5 and Fig. 6, which are just rotated and translated relative to each other. We use the value $\nu = (1.0 - \eta) \times 10^3$ to compare images, which is more convenient to use for η values which are close to 1.0. The greater ν is, the worse the correlation is. In the table below we show ν for $\phi = 0, \ldots, 5$.

ϕ	0	1	2	3	4	5
ν	0.016	0.017	0.017	0.017	0.016	0.016

We see that values are very close for different ϕ; thus we may use ν for any one of the ϕ values. The values are small which indicates very strong correlation.

If we compare Fig. 6 with a not-quite-the-same rotated and translated image (taken from an image after application of direct and inverse Fourier transform), we get a value of $\nu = 0.107$ which indicates a moderate correlation. The time to compute

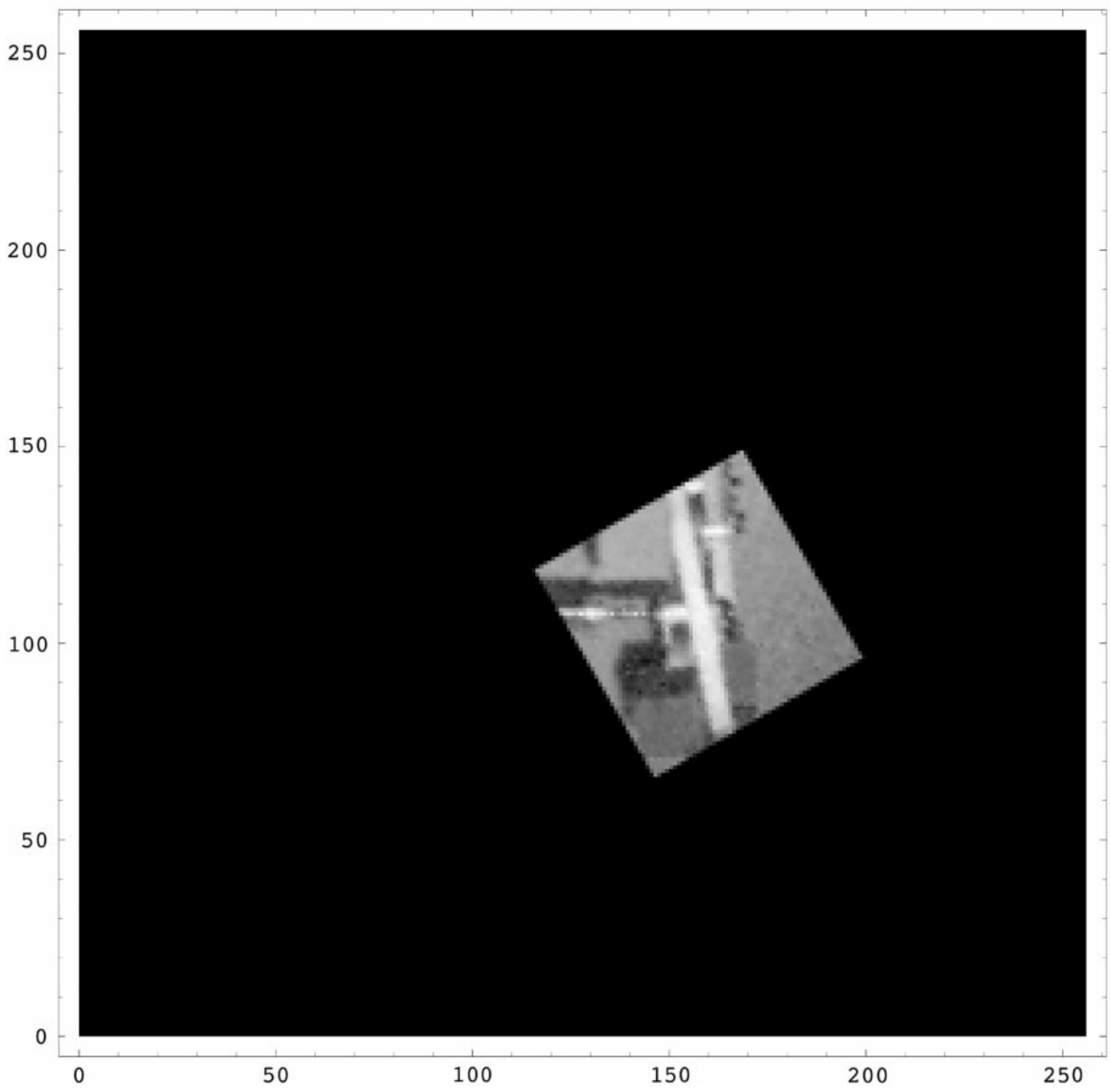

FIG. 6. A rotated and translated version of the image in Fig. 5.

direct Fourier transforms and correlations ν was around 3 s on a 250 MHz workstation.

8. CONCLUSIONS

In this paper we use techniques from noncommutative harmonic analysis to formulate problems in template matching and construction of image invariants. The main contribution is to illustrate that problems in image understanding can be cleanly formulated using mathematical techniques which are not standard tools in the community. Numerical examples are provided to demonstrate the techniques.

APPENDIX

In this appendix we review the essentials of noncommutative harmonic analysis which is the generalization of Fourier analysis to functions on groups. Much of the review material presented here may be found in [13, 14, 16, 25].

Recall that the Fourier transform pair for a suitable scalar function, $f(x)$, for $x \in \mathbb{R}$ is defined as

$$\hat{f}(p) = \int_{-\infty}^{\infty} f(x)u(-x, p)\,dx,$$
$$f(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \hat{f}(p)u(x, p)\,dp, \tag{A.1}$$

where $u(x, p) = e^{ipx}$. Note that $u(x + y, p) = e^{ip(x+y)} = e^{ipx} \cdot e^{ipy} = u(x, p)u(y, p)$. This is an example of a *group homomorphism*. In general, a homomorphism is a mapping between two groups $h : (G, \circ) \to (H, \hat{\circ})$ such that $h(g_1 \circ g_2) = h(g_1) \hat{\circ} h(g_2)$. In particular, the function $u(\cdot, p)$ maps $(\mathbb{R}, +) \to (U, \cdot)$ for each $\omega \in \mathbb{R}$, where U is the set of complex numbers with unit modulus, and $\cdot$ is scalar multiplication.

The *convolution theorem* for functions on the real line states that $(f_1(x) * f_2(x)) = \hat{f}_1(p)\hat{f}_2(p)$. This is a direct result of the facts that

$$u(-(x + y), p) = u(-x, p)u(-y, p)$$

and integration on the real line is translation invariant.

Noncommutative harmonic analysis extends the concept of Fourier transform and convolution to functions on groups. At the core of this area of mathematics is the enumeration of functions analogous to $u(x, p)$. Unlike the Abelian case where such functions are scalars, in the noncommutative context these functions are matrices called irreducible unitary representations (IURs).

A *representation* of a group G is a homomorphism $T : G \to T(G) \subset GL(V)$. V is a vector space called the representation space, and $GL(V)$ is the group of all invertible linear transformations of V onto itself. $T(g)$ for $g \in G$ is expressed in a given basis of V as an invertible matrix, and

$$T(g_1 \circ g_2) = T(g_1)T(g_2), \quad T(g^{-1}) = T^{-1}(g),$$
$$T(e) = 1 \in GL(V).$$

Representations that can be expressed as unitary matrices ($U^{-1} = U^{\dagger}$) in an orthonormal basis of V are called unitary representations. Irreducible representations are those which cannot be block-diagonalized. I.e., they are the "smallest" representations and cannot be further reduced. The function $u(x, p)$ is an example of a one-dimensional (and hence irreducible) unitary representation.

For a general unimodular group (i.e., a group which possesses a left and right invariant volume measure), the Fourier transform of a suitable function $f(g)$ is defined as

$$\hat{f}(\rho) = \mathcal{F}(f(g)) = \int_G f(g)U(g^{-1}; \rho)\, d\mu(g),$$

where $d\mu(g)$ is a left–right invariant volume measure on G and ρ is a dual (frequency-like) parameter which enumerates all the different IURs of the group. The parameter ρ could be a scalar, vector, or other quantity, depending on the group under consideration. The inverse Fourier transform

$$f(g) = \int_{\hat{G}} \text{trace}(\hat{f}(\rho)U(g; \rho))\, d\nu(\rho)$$

reconstructs the function from its spectrum (collection of Fourier transforms), where $d\nu(\rho)$ is an appropriately defined measure on the dual space of the group, $\hat{G}$.[4]

Because of the homomorphism property $U(g_1 \circ g_2; \rho) = U(g_1; \rho)U(g_2; \rho)$, the convolution theorem

$$\mathcal{F}(f_1 * f_2) = \hat{f}_2(\rho)\hat{f}_1(\rho)$$

holds, where convolution is defined as [13, 16, 14]

$$(f_1 * f_2)(g) = \int_G f_1(h) f_2(h^{-1} \circ g)\, d\mu(h).$$

The problem then becomes the enumeration of all inequivalent IURs for a given group.

A.1. Unitary Representations of SE(2)

A unitary representations of $SE(2)$ is defined by the unitary operator

$$U(g; p)\tilde{f}(\mathbf{x}) = e^{ip(\mathbf{b}\cdot\mathbf{x})}\tilde{f}(R^{\mathrm{T}}\mathbf{x}) \tag{A.2}$$

for each $g = (R, \mathbf{b}) \in SE(2)$. The form of this operator results from the semi-direct product structure of the group $SE(2)$; e^z is the scalar exponential function, $\rho = p \in \mathbb{R}^+$, $i = \sqrt{-1}$, and $\mathbf{x} \cdot \mathbf{y} = x_1y_1 + x_2y_2$. The vector $\mathbf{x}$ is a unit vector ($\mathbf{x} \cdot \mathbf{x} = 1$), and $\tilde{f}(\cdot) \in \mathcal{L}^2(S^1)$, where S^1 is the unit circle.

Since only one angle is required to parameterize a vector on the unit circle, $\mathbf{x} = (\cos\psi, \sin\psi)^{\mathrm{T}}$, and $\tilde{f}(\mathbf{x}) = \tilde{f}(\cos\psi, \sin\psi) \equiv f(\psi)$. Henceforth we will not distinguish between $\tilde{f}$ and f.

By definition, group representations observe the homomorphism property, which in this case is seen as

$$\begin{aligned} U(g_1; p)U(g_2; p)f(\mathbf{x}) &= U(g_1; p)(U(g_2; p)f(\mathbf{x})) \\ &= U(g_1; p)\big(e^{ip(\mathbf{b}_2\cdot\mathbf{x})} f\big(R_2^{\mathrm{T}}\mathbf{x}\big)\big) \\ &= e^{ip(\mathbf{b}_1\cdot\mathbf{x})}e^{ip(\mathbf{b}_2\cdot R_1^{\mathrm{T}}\mathbf{x})} f\big(R_2^{\mathrm{T}}R_1^{\mathrm{T}}\mathbf{x}\big) \\ &= e^{ip(\mathbf{b}_1 + R_1\mathbf{b}_2)\cdot\mathbf{x}} f((R_1R_2)^{\mathrm{T}}\mathbf{x}) \\ &= U(g_1 \circ g_2; p)f(\mathbf{x}). \end{aligned}$$

Any function $f(\psi) \in \mathcal{L}^2(S^1)$ can be expressed as a weighted sum of orthonormal basis functions as $f(\psi) = \sum_n a_n e^{in\psi}$. Likewise, the matrix elements of the operator $U(g; p)$ are expressed in this basis as

$$U_{mn}(g, p) = (e^{im\psi}, U(g; p)e^{in\psi}) \quad \forall m, n \in \mathbf{Z},$$

[4] It is worth noting that the measures $d\mu$ and $d\nu$ have not been determined for all unimodular groups, although they are well known for both the Euclidean group and the discrete motion group, which is all that is important in the context of this paper.

where the inner product $(\cdot, \cdot)$ is defined as

$$(f_1, f_2) = \frac{1}{2\pi}\int_0^{2\pi} f_1(\psi)\overline{f_2(\psi)}\, d\psi.$$

It is easy to see that $(U(g; p) f_1, U(g; p) f_2) = (f_1, f_2)$, and that $U(g; p)$ is therefore unitary with respect to this inner product.

The matrix with elements U_{mn} is "infinite dimensional." Furthermore, the matrix of a unitary operator expressed in an orthonormal basis is a unitary matrix, which means $U^{-1}_{nm} = \overline{u_{mn}}$.

A number of works including [22] have shown that the matrix elements of this representation are given by

$$u_{mn}(g(r, \phi, \theta), p) = i^{n-m} e^{-i[n\theta+(m-n)\phi]} J_{n-m}(pr), \quad \text{(A.3)}$$

where $J_\nu(x)$ is the νth order Bessel function and $g(r, \phi, \theta)$ is an element of $SE(2)$, where the translational part is parameterized in polar coordinates (r, ϕ).

A.2. *The Discrete Motion Group*

In order to get the IURs of the discrete motion group we restrict possible orientation angles to the values from C_N and choose the appropriate pulse orthonormal basis functions to compute the representation matrices using property (A.2).

We choose a pulse orthonormal basis $f_{N,n}(\mathbf{u})$ on S; i.e., we subdivide the circle into identical segments F_n and choose the f-functions to satisfy the orthonormality relations

$$\frac{1}{2\pi}\int_S f_{N,n}(\mathbf{u}) f_{N,m}(\mathbf{u})\, d\theta = \delta_{nm}.$$

We choose the orthonormal functions as

$$f_{N,n}(\mathbf{p}) = \begin{cases} (N)^{1/2} & \text{if } \mathbf{u} \in F_n, \\ 0 & \text{otherwise;} \end{cases}$$

$n = 0, \ldots, N-1$ enumerates different segments. We denote these pulse functions as δ-like functions $f_{N,n}(\mathbf{u}) = (1/N)^{1/2}$ $\delta_N(\mathbf{u}, \mathbf{u_n})$, where $\mathbf{u}_n$ is the vector to the center of the F_n segment.

The matrix elements in this basis are

$$U_{mn}(A, \mathbf{r}; p) = \frac{1}{2\pi}\int_S f_{N,m}(\mathbf{u}) e^{ip\mathbf{u}\cdot\mathbf{r}} f_{N,n}(A^{-1}\mathbf{u})\, d\theta. \quad \text{(A.4)}$$

It may be shown that this integral may be approximated as

$$U_{mn}(A_j, \mathbf{r}; p) = e^{ip\mathbf{u}_m\cdot\mathbf{r}} \delta_{A_j^{-1}\mathbf{u}_m, \mathbf{u}_n}, \quad \text{(A.5)}$$

where $\delta_{A_j^{-1}\mathbf{u}_m,\mathbf{u}_n} = \delta_{m-j,n}$ in this case.

The matrix elements (A.5) are exact expressions for the matrix elements of the unitary representations of the discrete motion group. The set of matrix elements (A.5) is, however, incomplete. This means, that the direct and inverse Fourier transforms, defined using these matrix elements, would reproduce the original function with $O(1/N)$ error; i.e.,

$$\mathcal{F}^{-1}(\mathcal{F}(f(A_i, \mathbf{r})) = f(A_i, \mathbf{r})(1 + O(1/N)).$$

The reason for this is that summing through all possible segments cannot replace integration over all possible angles on the circle. It is also clear that the additional continuous parameter which enumerates possible angles inside each segment on the circle must give the complete set of the matrix elements.

Thus, the matrix elements must be modified as

$$U_{mn}(A_j, \mathbf{r}; p, \phi) = e^{ip\mathbf{u}_m^\phi\cdot\mathbf{r}} \delta_{A_j^{-1}\mathbf{u}_m, \mathbf{u}_n}, \quad \text{(A.6)}$$

where $\mathbf{u}_k^\phi$ denotes the vector to the angle $\theta = \phi + 2\pi k/N$ on the unit circle on the interval $[2\pi k/N, 2\pi(k+1)/N]$, $k = 0, \ldots, N-1$ (ϕ measures the angle on this segment). The vectors $\mathbf{u}_m^\phi$ are illustrated in Fig. 7. The Fourier parameter in this case is the pair $\rho = (p, \phi)$.

This is expression (7) in the text.

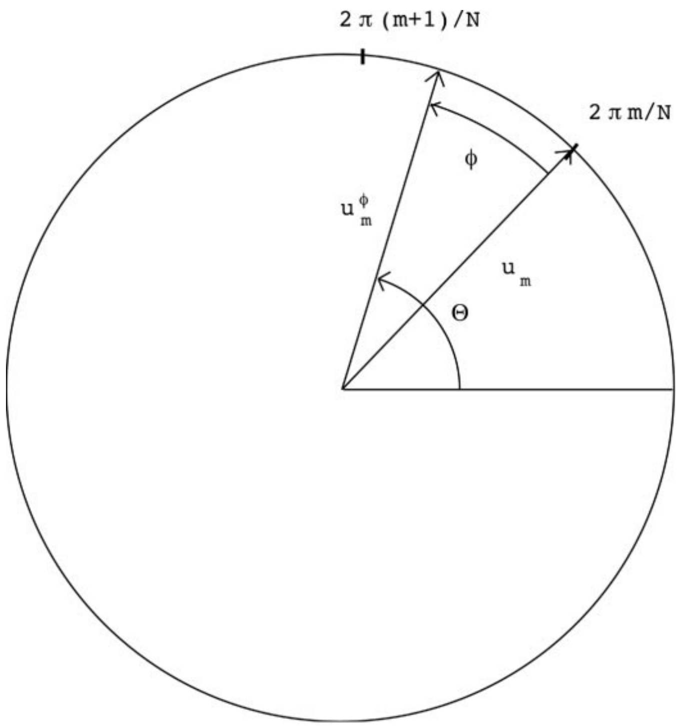

FIG. 7. Illustration for vectors $\mathbf{u}_m^\phi$ in the matrix elements of IURs of the discrete motion group.

REFERENCES

1. B. Jahne, *Spatio-Temporal Image Processing: Theory and Scientific Applications*, Springer-Verlag, Berlin, 1993.
2. G. L. Turin, An introduction to matched filters, *IRE Trans. Inf. Theory* **IT-6**, 1960, 311–329.
3. M. A. Karim and A. A. S. Awwal, *Optical Computing An Introduction*, Wiley, New York, 1992.
4. H. H. Arsenault and Y. N. Hsu, Chalasinsk–Macukow, rotation-invariant pattern recognition, *Opt. Eng.* **23**, 1984, 705–709.
5. A. B. Bhatia and E. Wolf, On the circle polynomials of Zernike and related orthogonal sets, *Proc. Cambridge Philos. Soc.* **50**, 1954, 40–48.
6. Y. S. Abu-Mostafa and D. Psaltis, Recognition aspects of moment invariants, *IEEE Trans. Pattern Anal. Mach. Intell.* **PAMI-6**, 1984, 698.
7. M. Antonini, M. Barlaud, P. Mathieu, and I. Daubechies, Image coding using wavelet transform, *IEEE Transactions on Image Processing* **40**(2), 1992, 205–220.
8. J.-P. Leduc, Spatio-temporal wavelet transforms for digital signal analysis, *Signal Processing* **60**, 1997, 23–41.
9. R. Murenzi, Wavelet transforms associated to the *n*-dimensional Euclidean group with dilations: Signals in more than one dimension, in *Wavelets: Time-Frequency Methods and Phase Space* (J. M. Combes, A. Grossmann, and Ph. Tchamitchian, Eds.), pp. 239–246.
10. J. Segman and Y. Zeevi, Image analysis by wavelet-type transforms: Group theoretical approach, *Journal of Mathematical Imaging and Vision* **3**, 1993, 51–77. [“Estimation with a Pattern Recognition (ICPR’86), Washington DC, 1986”]
11. J. P. Gauthier, G. Bornard, and M. Sibermann, Motion and pattern analysis: Harmonic analysis on motion groups and their homogeneous spaces, *IEEE Trans. Syst. Man Cybern*. **21**, 1991, 159–172.
12. R. Lenz, *Group Theoretical Methods in Image Processing*, Lecture Notes in Computer Science, Springer-Verlag, Berlin/Heidelberg/New York, 1990.
13. G. S. Chirikjian and I. Ebert-Uphoff, Numerical convolution on the Euclidean group with applications to workspace generation, *IEEE Trans. Robotics Automation* **14**(1), 1998, 123–136.
14. G. B. Folland, *A Course in Abstract Harmonic Analysis*, CRC Press, Boca Raton, FL, 1995.
15. K. Kanatani, *Group-Theoretical Methods in Image Understanding*, Springer-Verlag, Berlin/Heidelberg/New York, 1990.
16. G. Chirikjian, Fredholm integral equations on the Euclidean motions group, *Inverse Problems* **12**, 1996, 579–599.
17. J. W. Cooley and J. Tukey, An algorithm for the machine calculation of complex Fourier series, *Math. Comput.* **19**, 1965, 297–301.
18. D. F. Elliott and K. R. Rao, *Fast Transforms: Algorithms, Analyses, Applications*, Academic Press, New York, London, 1982.
19. H. Choi and D. C. Munson, Direct-Fourier reconstruction in tomography and synthetic aperture radar, *Int. J. of Imaging Systems and Technology* **9**, 1998, 1–13.
20. H. Stark, J. W. Woods, I. Paul, and R. Hingorani, Direct Fourier reconstruction in computed tomography, *IEEE Trans. Acoustics, Speech, Signal Processing* **ASSP-29**, 1981, 237–245.
21. A. Kyatkin and G. Chirikjian, Regularized solutions of a nonlinear convolution equation on the Euclidean group, *Acta Appl. Math.* **53**, 1986, 89–123.
22. N. J. Vilenkin, Bessel functions and representations of the group of Euclidean motions, *Uspehi Mat. Nauk.* **11**, 1956, 69–112. [Russian]
23. J. Canny. A computational approach to edge detection, *IEEE Trans. Pattern Anal. Mach. Intell.* **8**, 1986, 679–698.
24. H. Dym and H. McKean, *Fourier Series and Integrals*, Academic Press, New York, 1972.
25. M. Sugiura, *Unitary Representations and Harmonic Analysis*, 2nd ed., Elsevier Science, Amsterdam, 1990.
26. D. Gurarie, *Symmetry and Laplacians. Introduction to Harmonic Analysis, Group Representations and Applications*, Elsevier Science, 1992.