CrossMark

# Fourier Descriptors Based on the Structure of the Human Primary Visual Cortex with Applications to Object Recognition

Amine Bohi[1] · Dario Prandi[2] · Vincente Guis[1] · Frédéric Bouchara[1] · Jean-Paul Gauthier[1]

**Abstract** In this paper, we propose a supervised object recognition method using new global features and inspired by the model of the human primary visual cortex V1 as the semidiscrete roto-translation group $SE(2,N) = \mathbb{Z}_N \rtimes \mathbb{R}^2$. The proposed technique is based on generalized Fourier descriptors on the latter group, which are invariant to natural geometric transformations (rotations, translations). These descriptors are then used to feed an SVM classifier. We have tested our method against the COIL-100 image database and the ORL face database, and compared it with other techniques based on traditional descriptors, global and local. The obtained results have shown that our approach looks extremely efficient and stable to noise, in presence of which it outperforms the other techniques analyzed in the paper.

**Keywords** Descriptor · Fourier transform · Hexagonal grid · Geometric transformations · Support vector machine · Object recognition

✉ Amine Bohi
amine.bohi@lsis.org

Dario Prandi
prandi@ceremade.dauphine.fr

Vincente Guis
guis@univ-tln.fr

Frédéric Bouchara
bouchara@univ-tln.fr

Jean-Paul Gauthier
gauthier@univ-tln.fr

[1] LSIS Laboratory, University of South Toulon - Var, B.P. 20132 83957, La Garde Cedex, France

[2] CEREMADE Laboratory, University of Paris-Dauphine, Place du Maréchal De Lattre De Tassigny, 75775 Paris Cedex 16, France

## 1 Introduction

Object recognition is a fundamental problem in computer vision and keeps attracting more and more attention nowadays. Its concepts have been applied in multiple fields, such as manufacturing, surveillance system, optical character recognition, face recognition, etc.

Almost every object recognition algorithm proposed in the literature is based on the computation of certain features of the image, which allow to characterize the object depicted and to discriminate it from others. In particular, since objects can appear at different locations and with different sizes, it is desirable for such features to be invariant by translation, rotation and scale. These invariant features can be global, i.e., computed taking into account the whole image, or local, i.e., computed considering only neighborhoods of key-points in the image.

In this paper we focus on *Fourier descriptors*, an important class of global invariant features used since the seventies [19, 40] based on algebraic properties of the Fourier transform. In particular, inspired by some neurophysiological facts on the structure of the human primary visual cortex, we extend this theory to define features invariant to translation and rotation and we apply them for invariant object recognition in SVM context. These results are then compared with those obtained with another important class of global invariant features, the *moment invariants* (see Appendix 1), used e.g., in [10,34,35], and with two different local invariants. For more information on object recognition via local features we refer to [2,12,25, 27,28,30].

Our choice of a global approach is motivated by the better results obtained by these methods in presence of noise, luminance changes and other different alterations, with respect to algorithms based on local invariant features [7]. Indeed, under these conditions, key-points detectors used in the local

approach produce key-points that are not relevant for the object recognition.

In the following, we briefly introduce the theory of Fourier descriptors, before discussing the framework used in this paper and our contributions.

## 1.1 Fourier Descriptors

The basic idea behind Fourier descriptors is that the action of an abelian locally compact group $\mathbb{G}$ on functions in $L^2(\mathbb{G})$ is much easier to treat at the level of their Fourier transforms. In the specific case of images, $f, g \in L^2(\mathbb{R}^2)$, this is expressed by the well-known equivalence for the translation of $a \in \mathbb{R}^2$:

$$f(x) = g(x - a) \quad \forall x \in \mathbb{R}^2$$
$$\iff \hat{f}(\lambda) = e^{i\langle a,\lambda\rangle}\hat{g}(\lambda) \quad \forall \lambda \in \mathbb{R}^2, \quad (1)$$

where the Fourier transform is defined[1] by

$$\hat{f}(\lambda) = \int_{\mathbb{R}^2} f(x)\, e^{-i\langle\lambda,x\rangle}\, dx, \quad \forall \lambda \in \mathbb{R}^2.$$

In this setting, Fourier descriptors are quantities associated with functions of $L^2(\mathbb{R}^2)$ that can be easily computed starting from their Fourier representation and that are invariant under the action of translations. Ideally, a Fourier descriptor should be **complete**, meaning that for any couple of images $f, g \in L^2(\mathbb{R}^2)$ the equality of the Fourier descriptor is equivalent to the equality of $f$ and $g$ up to translations. Indeed, the lack of completeness could yield to problems in applications, notably to false positives in the classification.

However, a result as strong as completeness is usually out of reach and unnecessary for practical applications. In this case, one looks for Fourier descriptors that are at least **weakly complete**, meaning that they are complete on a sufficiently big subset of $L^2(\mathbb{R}^2)$, usually either open and dense or at least residual, i.e., the intersection of countably many open and dense sets. This guarantees that the Fourier descriptor will correctly classify a sufficiently large class of images.

Various Fourier descriptors have been defined in the literature [19,24,26,38,40]. In this work we are mainly interested in the following two, whose invariance w.r.t. translations can be checked via (1).

– **Power-spectrum:** the quantity $\mathrm{PS}_f(\lambda) := |\hat{f}(\lambda)|^2$ for $\lambda \in \mathbb{R}^2$, which is the Fourier transform of the auto-correlation function

$$a_f(x) := \int_{\mathbb{R}^2} \overline{f(y)}\, f(y + x)\, dy.$$

It is easy to show that the power-spectrum is not weakly complete, and indeed it is used in texture synthesis to identify the translation invariant Gaussian distribution of textures [18].

– **Bispectrum:** an extension of the power-spectrum, it is the quantity $\mathrm{BS}_f(\lambda_1, \lambda_2) := \hat{f}(\lambda_1)\,\hat{f}(\lambda_2)\,\overline{\hat{f}(\lambda_1 + \lambda_2)}$, or equivalently the Fourier transform of the triple correlation, defined as

$$a_{3,f}(x_1, x_2) := \int_{\mathbb{R}^2} \overline{f(y)}\, f(y + x_1)\, f(y + x_2)\, dy.$$

These descriptors are complete on compactly supported functions of $L^2(\mathbb{R}^2)$ and are well established in statistical signal processing. See e.g., [14], where they are applied to sound texture recognition.

These two Fourier descriptors can be easily generalized to functions on $L^2(\mathbb{G})$ of a locally compact abelian group $\mathbb{G}$ to obtain invariants under the action of $\mathbb{G}$. This can be applied, for example, to 2D shape recognition. However, when working with images, these descriptors are unsatisfying. Indeed, they are invariant only under translations, and so cannot be used to classify images under the action of rotations.

## 1.2 Framework of the Paper

In this paper, following a line of research started in [38], we present a theoretical framework that allows us to build generalized Fourier descriptors which are invariant w.r.t. (semidiscrete) roto-translations of images. We exploit the following two facts:

– It is possible to define a natural generalization of the power-spectrum and the bispectrum on non-commutative groups, as it has been done in [24,38].
– Contributions of some of the authors to a fairly recent model of the human primary visual cortex V1 [4,5] have shown that the latter can be modeled as the semidiscrete group of roto-translations $SE(2, N) = \mathbb{Z}_N \rtimes \mathbb{R}^2$. In this model, cortical stimuli are functions in $L^2(SE(2, N))$, w.r.t. the Haar measure of $SE(2, N)$, and images from the visual plane are lifted to cortical stimuli via a natural injective and left-invariant lift operation $\mathcal{L} : L^2(\mathbb{R}^2) \to L^2(SE(2, N))$. Such lift is defined as the wavelet transform w.r.t. a mother wavelet $\Psi$, see Sect. 2.

From these facts, a natural pipeline for invariant object recognition is the following:

1. Given an image $f \in L^2(\mathbb{R}^2)$ lift it to a cortical stimulus $\mathcal{L}f \in L^2(SE(2, N))$.

---

[1] Here we use a non-unitary definition of the Fourier transform for future convenience in computations.

2. Compute the generalized Fourier descriptors of $\mathcal{L}f$ on the non-commutative group $SE(2, N)$.

3. If the lift of another image $g \in L^2(\mathbb{R}^2)$ has the same Fourier descriptors as $\mathcal{L}f$, deduce that $\mathcal{L}f \approx \mathcal{L}g$ up to the action of $SE(2, N)$.

4. Thanks to the left-invariance and injectivity of the lift $\mathcal{L}$, obtain that also $f \approx g$ up to the action of $SE(2, N)$.

This pipeline was already investigated in [38], where the authors considered a non-left-invariant lift, the *cyclic lift*. For this lift they then proved a weak completeness result of the generalized bispectrum for images, represented as $L^2(\mathbb{R}^2)$ functions with support inside a fixed compact set.

In this paper, we consider the same question for left-invariant lifts, where the situation turns out to be more complicated. In particular, as explained in the following section, to ensure the weak completeness we are led to consider "stronger" invariants than the generalized bispectrum. However, as observed in Remark 4, the actual computation of these stronger invariants on lifted images requires $N$ times less computational time and space w.r.t. the computation of the generalized bispectrum of cyclically lifted images.

### 1.3 Contributions of the Paper

Let $K \subset \mathbb{R}^2$ be any compact set, representing the size of the images under consideration. According to the pipeline for object recognition introduced above, the weak completeness of the generalized Fourier descriptors on images can be proved in two steps:

1. Prove the completeness of the generalized Fourier descriptors on some residual set $\mathcal{G} \subset L^2(\mathbb{Z}_N \times K)$ of cortical stimuli;

2. Prove that for some residual set $\mathcal{R} \subset L^2(\mathbb{R}^2)$ of images with support in $K$ we have $\mathcal{L}(\mathcal{R}) \subset \mathcal{G}$.

The first point is addressed in Theorem 1, where is identified an open and dense set $\mathcal{G} \subset L^2(\mathbb{Z}_N \times K)$ on which the combination of the generalized power-spectrum and bispectrum holds. This generalizes the result in [38], where the same result was proved for a residual subset of the range of the cyclic lift.

Unfortunately, it turns out that for this set $\mathcal{G}$ and a left-invariant lift $\mathcal{L}$ there is no hope of finding a set $\mathcal{R} \subset L^2(\mathbb{R}^2)$ satisfying the second point above. We are then led to introduce stronger Fourier descriptors, the rotational power-spectrum and bispectrum, which are invariant only w.r.t. rotations. To solve this problem we preprocess images by centering them at their barycenter, a procedure that is essential also in [38]. Theorem 2 then shows that the resulting invariants are complete for an open and dense set of functions in $L^2(K)$, for any compact $K \subset \mathbb{R}^2$. The proof of this com-

pleteness requires fine technical tools from harmonic analysis and the theory of circulant operators, and for this reason we only present a sketch of it, evidencing the technical difficulties. A complete proof will be presented in a forthcoming paper by the second and last authors.

Finally, in Theorem 3 we show that, under mild assumptions on the mother wavelet $\Psi$, to check the equality of all these Fourier descriptors it is sufficient to compute simple quantities computed from the 2D Fourier transform of the image. This allows for an efficient implementation on regular hexagonal grids. After using these descriptors to feed an SVM-based classifier, we compare their performances with those of Hu and Zernike moments, the Fourier–Mellin transform and some well-known local descriptors. To this purpose, we test them on two large databases: the COIL-100[2] object recognition database, composed of 7200 objects presenting rotation and scale changes [31], and the ORL[3] face database, on which different human faces are subject to several kinds of variations.

### 1.4 Structure of the Paper

The remainder of the paper is organized as follows. In Sect. 2, we present the features of a mathematical model of the primary visual cortex $V1$ that are essential to our approach. In Sect. 3, we introduce some generalities on the Fourier Transform on the semidiscrete group of roto-translation $SE(2, N)$. In Sect. 4, we describe the natural generalization of the power-spectrum and the bispectrum on $\mathbb{R}^2$ to $SE(2, N)$. We then prove the weak completeness result (Theorem 1) and show that under the chosen lift operator this does not imply weak completeness for images. Finally, we introduce the rotational power-spectrum and bispectrum and sketch the proof of the corresponding weak completeness result (Theorem 2) for images. We end this section with some results on the practical computation of these descriptors. In Sect. 5, we illustrate some numerical results where these descriptors are compared with those obtained via global descriptors such as Zernike moments, Hu moments, Fourier–Mellin transform, and local ones like the SIFT and HoG descriptors. Finally, we conclude with some practical suggestions in Sect. 6.

## 2 A Mathematical Model of the Primary Visual Cortex V1

As mentioned in the introduction, the main novelty of our approach is its connection with a fairly recent model of the human primary visual cortex V1 due to Petitot and Citti-Sarti [11,32] and our recent contributions [3–5,33]. The theory

---

of orientation scores introduced in [15,16] is also strongly connected with this work, in particular for its exploitation of left-invariant lift operators. We also mention [37], where image invariants based on the structure of the roto-translation group $SE(2)$ are introduced for textures. In this section, we present the features of this model that are essential to our approach.

Since it is well-known [23] that neurons in V1 are sensitive not only to positions in the visual field, but also to local orientations and that it is reasonable to assume these orientations to be finite, in [4] V1 has been modeled as the semidiscrete group of roto-translations $SE(2, N) = \mathbb{Z}_N \rtimes \mathbb{R}^2$ for some even $N \in \mathbb{N}$. Letting $R_k$ be the rotation of $2\pi/k$, the (non-commutative) group operation of $SE(2, N)$ is

$$(x, k)(y, r) = (x + R_k y, k + r).$$

Here, we are implicitly identifying $k + r$ with $k + r \mod N$.

Visual stimuli $f \in L^2(\mathbb{R}^2)$ are assumed to be lifted to activation patterns in $L^2(SE(2, N))$ by a lift operator $\mathcal{L} : L^2(\mathbb{R}^2) \to L^2(SE(2, N))$. Motivated by neurophysiological evidence, we then assume that

(H)  The lift operator $\mathcal{L}$ is linear and is defined as

$$\mathcal{L}f(x, k) := \int_{\mathbb{R}^2} f(y)\bar{\Psi}(R_{-k}(y - x)) \, dy, \qquad (2)$$

for a given mother wavelet $\Psi \in L^2(\mathbb{R}^2)$ such that $\mathcal{L}$ is injective and bounded.

*Remark 1* This assumption means that the lift operator under consideration is the wavelet transform w.r.t. $\Psi$ (see e.g., [17]). The fact that $\mathcal{L}$ be injective and bounded is then equivalent to the fact that the mother wavelet $\Psi$ is *weakly admissible*, i.e., is such that the map

$$\lambda \in \mathbb{R}^2 \mapsto \sum_{k \in \mathbb{Z}_N} |\hat{\Psi}(R_{-k}\lambda)|^2$$

is strictly positive and essentially bounded.

As a consequence of the above assumption, the lift operation $\mathcal{L}$ is left-invariant w.r.t. to the action of $SE(2, N)$. Namely,

$$\Lambda(x, k) \circ \mathcal{L} = \mathcal{L} \circ \pi(x, k). \qquad (3)$$

Here $\Lambda$ and $\pi$ are the actions of $SE(2, N)$ on $L^2(SE(2, N))$ and $L^2(\mathbb{R}^2)$, respectively. That is,

$$[\Lambda(x, k)\varphi](y, r) = \varphi\left((x, k)^{-1}(y, r)\right)$$
$$= \varphi(R_{-k}(y - x), r - k),$$
$$[\pi(x, k)f](y) = f\left((x, k)^{-1}y\right) = f(R_{-k}(y - x)).$$

Formula (3) can be seen as a semidiscrete version of the shift-twist symmetry [6].

The main observation for our purposes is that (3) means that two images $f$ and $g \in L^2(\mathbb{R}^2)$ can be deduced via roto-translation (i.e., $f = \pi(x, k)g$ for some $(x, k) \in SE(2, N)$) if and only if their lifts can be deduced via $\Lambda(x, k)$.

## 3 Preliminaries on Non-commutative Harmonic Analysis

In this section we introduce some generalities on the (non-commutative) Fourier transform on $SE(2, N)$, an essential tool to define and compute the Fourier descriptors we are interested in. We refer to [1,20] for a general introduction to the topic.

Since $SE(2, N)$ is a non-commutative unimodular group, the Fourier transform of $\varphi \in L^2(SE(2, N))$ is an operator associating to each (continuous) irreducible unitary representations $T^\lambda$ of $SE(2, N)$ some Hilbert–Schmidt operator on the Hilbert space where $T^\lambda$ acts. Here, $\lambda$ is an index taking values in the dual object of $SE(2, N)$, which is denoted by $\widehat{SE(2, N)}$ and is the set of equivalence classes of irreducible unitary representations.

The set of irreducible representations of a semi-direct product group can be obtained via Mackey's machinery (see e.g., [1, Ch. 17.1, Theorems 4 and 5]). Accordingly, $\widehat{SE(2, N)}$ is parametrized by the orbits of the (contragredient) action of rotations $\{R_k\}_{k \in \mathbb{Z}_N}$ on $\mathbb{R}^2$, i.e., by the slice $\mathcal{S} \subset \mathbb{R}^2$ which in polar coordinates is $(0, +\infty) \times [0, 2\pi/N)$. Additionally, corresponding to the origin, there are the characters of $\mathbb{Z}_N$. Namely, to each $\lambda \in \mathcal{S}$ corresponds the representation $T^\lambda$ acting on $\mathbb{C}^N$ via

$$T^\lambda(x, k)v = \text{diag}_h(e^{i\langle \lambda, R_h x \rangle}) \circ S^k v$$
$$= \left(e^{i\langle \lambda, R_h x \rangle} v_{h+k}\right)_{h=0}^{N-1}, \qquad (4)$$

where we denoted by $\text{diag}_h v_h$ the diagonal matrix of diagonal $v \in \mathbb{C}^N$ and by $S$ the shift operator $(Sv)_j = v_{j+1}$, so that $(S^k v)_j = v_{j+k}$. On the other hand, to each $k \in \mathbb{Z}_N$ corresponds the representation on $\mathbb{C}$ given by $z \mapsto e^{i\frac{2\pi k}{N}}z$. Since it is possible to show that to invert the Fourier transform it is enough to consider only the representations parametrized by $\mathcal{S}$, we will henceforth ignore the $\mathbb{Z}_N$ part of the dual.

Finally, the matrix-valued Fourier coefficient of a function $\varphi \in L^2(SE(2, N)) \cap L^1(SE(2, N))$ for $\lambda \in \widehat{SE(2, N)}$ is

$$\hat{\varphi}(T^\lambda) = \int_{SE(2,N)} \varphi(a)\, T^\lambda(a^{-1})\, da, \tag{5}$$

where $da$ is the Haar measure[4] of $SE(2,N)$. This is essentially the same formula for the Fourier transform on $\mathbb{R}^2$, which is a scalar and is obtained using the representations $\lambda(x) = e^{i\langle x,\lambda\rangle}$ acting on $\mathbb{C}$.

Straightforward computations yield

$$\hat{\varphi}(T^\lambda)_{i,j} = \mathcal{F}(\varphi(\cdot, i-j))(R_{-j}\lambda), \tag{6}$$

where we let $\mathcal{F}$ denote the Fourier transform on $\mathbb{R}^2$.

As usual, the definition of the Fourier transform can be extended to the whole $L^2(SE(2,N))$ by density arguments. Then, there exists a unique measure on $\widehat{SE(2,N)}$, the Plancherel measure, supported on $\mathcal{S}$ where it coincides with the restriction of the Lebesgue measure of $\mathbb{R}^2$, such that the Fourier transform is an isometry between $L^2(SE(2,N))$ and $L^2(\widehat{SE(2,N)})$. In particular, the following inversion formula holds

$$\varphi(x,k) = \int_{\mathcal{S}} \mathrm{Tr}\left(\hat{\varphi}(T^\lambda) \circ T^\lambda(x,k)\right)\, d\lambda. $$

The fundamental property of the non-commutative Fourier transform, generalizing (1), is that for all $\varphi, \eta \in L^2(SE(2,N))$ and $a \in SE(2,N)$ it holds

$$\begin{aligned} \varphi(x,k) = [\Lambda(a)\eta](x,k) \quad \forall(x,k) \in SE(2,N) &\iff \\ \hat{\varphi}(T) = \hat{\eta}(T) \circ T^{-1}(a) \quad \forall T \in \widehat{SE(2,N)}. \end{aligned} \tag{7}$$

Namely, $\varphi, \eta$ can be deduced via the action of $SE(2,N)$ if and only if their Fourier transforms at a representation $T$ can be deduced via multiplication by $T(a)$.

*Remark 2* The fact that the Fourier transform in (5) be matrix-valued is a direct consequence of $SE(2,N)$ being a Moore group, that is, that all the $T^\lambda$ act on finite-dimensional spaces. This is not true for the roto-translation group $SE(2)$. As a consequence, the Fourier transform on $SE(2)$, takes values not in the finite dimensional space of complex $N \times N$ matrices, but in the infinite dimensional space of operators over $L^2(\mathbb{S}^1)$. This is indeed the main theoretical advantage of considering $SE(2,N)$.

### 3.1 Decomposition of Tensor Product Representations

Proofs of Sect. 4, will use a well-known fact on tensor product representations: the Induction–Reduction Theorem (see [1]). This theorem allows to decompose the tensor products of

representations $T^{\lambda_1} \otimes T^{\lambda_2}$, acting on $\mathbb{C}^N \otimes \mathbb{C}^N \cong \mathbb{C}^{N \times N}$, to an equivalent representation acting on $\bigoplus_{k \in \mathbb{Z}_N} \mathbb{C}^N$, which is a block-diagonal operator whose block elements are of the form $T^{\lambda_1 + R_k \lambda_2}$. Moreover, the linear transformation realizing the equivalence is explicit.

To avoid confusion, we will henceforth denote components of vectors $v \in \mathbb{C}^N$ as $v(0), \ldots, v(N-1)$, elements of $\bigoplus_{k \in \mathbb{Z}_N} \mathbb{C}^N$ as $(w_k)_{k \in \mathbb{Z}_N}$ where $w_k \in \mathbb{C}^N$, and the components of vectors $\mathbf{v} \in \mathbb{C}^N \otimes \mathbb{C}^N$ as $\mathbf{v}(k,h)$ for $k, h = 0, \ldots, N-1$. We also remark that linear operators $\mathcal{B}$ on $\bigoplus_{k \in \mathbb{Z}_N} \mathbb{C}^N$ can be decomposed as $\mathcal{B} = (\mathcal{B}^{k,h})_{k,h \in \mathbb{Z}_N}$, where each $\mathcal{B}^{k,h}$ is an $N \times N$ complex matrix. Namely, we have

$$\mathcal{B}(w_k)_{k \in \mathbb{Z}_N} = \left(\sum_{h=0}^{N-1} \mathcal{B}^{k,h} w_h\right)_{k \in \mathbb{Z}_N}. \tag{8}$$

Then, exploiting the commutation of the Fourier transform with equivalences of representation, the Induction–Reduction Theorem implies that for every $\varphi \in L^2(SE(2,N))$ and any $\lambda_1, \lambda_2 \in \mathcal{S}$ it holds

$$A \circ \hat{\varphi}(T^{\lambda_1} \otimes T^{\lambda_2}) \circ A^{-1} = \bigoplus_{k \in \mathbb{Z}_N} \hat{\varphi}(T^{\lambda_1 + R_k \lambda_2}). \tag{9}$$

Here, $A : \mathbb{C}^N \otimes \mathbb{C}^N \to \bigoplus_{k \in \mathbb{Z}_N} \mathbb{C}^N$ is given by

$$(A\mathbf{v})_k(h) = (A_k \mathbf{v})(h) = \mathbf{v}(h, h-k), \quad \forall \mathbf{v} \in \mathbb{C}^N \otimes \mathbb{C}^N. \tag{10}$$

## 4 Fourier Descriptors on $SE(2,N)$

In the following sections, we introduce and study the Fourier descriptors on the group $SE(2,N)$. As already mentioned, proving a general completeness result is essentially hopeless, and we will content ourselves to prove the weak completeness.

Let $K \subset \mathbb{R}^2$ be a compact set. In the following, we will be mainly concerned with functions that are compactly supported either in $K$ or in $K \times \mathbb{Z}_N \subset SE(2,N)$.

### 4.1 Generalized Fourier Descriptors

Following [38], the power-spectrum and the bispectrum on $\mathbb{R}^2$ can be generalized to $SE(2,N)$ as follows.

**Definition 1** The *generalized power-spectrum* and *bispectrum* of $\varphi \in L^2(SE(2,N))$ are the collections of matrices for any $\lambda, \lambda_1, \lambda_2 \in \mathcal{S}$,

$$\mathrm{PS}_\varphi(\lambda) := \hat{\varphi}(T^\lambda) \circ \hat{\varphi}(T^\lambda)^*$$
$$\mathrm{BS}_\varphi(\lambda_1, \lambda_2) := \hat{\varphi}(T^{\lambda_1}) \otimes \hat{\varphi}(T^{\lambda_2}) \circ \hat{\varphi}(T^{\lambda_1} \otimes T^{\lambda_2})^*.$$

---

The next result generalizes, with a simplified proof, the result presented in [38]. Let us mention that this result is indeed true in a more general setting, as it will be shown in a forthcoming paper by Prandi and Gauthier.

**Theorem 1** *Let $K \subset \mathbb{R}^2$ be a compact. The generalized power-spectrum and bispectrum are weakly complete on $L^2(\mathbb{Z}_N \times K)$. In particular, they discriminate on the open and dense set $\mathcal{G} \subset L^2(\mathbb{Z}_N \times K)$ of functions $\varphi$ supported in $\mathbb{Z}_N \times K$ and whose Fourier transform $\hat{\varphi}(T^\lambda)$ is invertible for an open and dense set of $\lambda$s. That is, $\varphi_1, \varphi_2 \in \mathcal{G}$ are such that $\mathrm{PS}_{\varphi_1} = \mathrm{PS}_{\varphi_2}$ and $\mathrm{BS}_{\varphi_1} = \mathrm{BS}_{\varphi_2}$ if and only if $\varphi_1 = \Lambda(x, k)\varphi_2$ for some $(x, k) \in SE(2, N)$.*

*Proof* The fact that $\mathcal{G}$ is open and dense is proved in Lemma 1 in Appendix 2. Let $\varphi, \eta \in \mathcal{G}$ be such that $PS_{\varphi_1} = PS_{\varphi_2}$ and $\mathrm{BS}_\varphi = \mathrm{BS}_\eta$. The equality of the generalized bispectrum implies that the set of $\lambda$s where $\hat{\varphi}(T^\lambda)$ and $\hat{\eta}(T^\lambda)$ fail to be invertible is the same. We will denote it by $I$ and let

$$U(T^\lambda) := \hat{\varphi}(T^\lambda)^{-1}\hat{\eta}(T^\lambda) \quad \forall \lambda \in I.$$

In order to complete the proof of the statement, we will prove that $U(T^\lambda)$ can be defined for all $\lambda$s in in $\mathbb{R}^2$ and, moreover, that $U(T^\lambda) = T^\lambda(a)$ for some $a \in SE(2, N)$. Indeed, by (7) this will readily imply that $\varphi = \Lambda(a)\eta$ as announced.

We claim that $U(T^\lambda)$ is unitary for all $\lambda \in I$. Indeed, by the equality of the generalized power-spectrum we have

$$U(T^\lambda)^*U(T^\lambda) = \eta(T^\lambda)^* \mathrm{PS}_\varphi(\lambda)\eta(T^\lambda) = \mathrm{I}.$$

Observe that the equality of the generalized bispectrum and the definition of $U$, imply that for all $\lambda_1, \lambda_2 \in I$ it holds

$$\mathrm{BS}_\varphi(\lambda_1, \lambda_2) = \hat{\varphi}(T^{\lambda_1}) \otimes \hat{\varphi}(T^{\lambda_2}) \circ U(T^{\lambda_1})$$
$$\otimes U(T^{\lambda_2}) \circ \hat{\eta}(T^{\lambda_1} \otimes T^{\lambda_2})^*.$$

By the invertibility of $\hat{\varphi}(T^{\lambda_1}) \otimes \hat{\varphi}(T^{\lambda_2})$ and the unitarity of $U$, this yields

$$\hat{\varphi}(T^{\lambda_1} \otimes T^{\lambda_2}) \circ U(T^{\lambda_1}) \otimes U(T^{\lambda_2}) = \hat{\eta}(T^{\lambda_1} \otimes T^{\lambda_2}). \quad (11)$$

The announced result is then a consequence of the following three facts, which are proved in Appendix 2.

1. Lemma 2: The function $\lambda \mapsto U(T^\lambda)$ is continuous on $I$.
2. Lemma 3: The function $\lambda \mapsto U(T^\lambda)$ can be extended to a continuous function on $\mathbb{R}^2$ for which (11) is still true.
3. Lemma 4: There exists $a \in SE(2, N)$ such that $U(T^\lambda) = T^\lambda(a)$. $\square$

An immediate corollary is the following.

**Corollary 1** *Let $\tilde{\mathcal{L}} : L^2(\mathbb{R}^2) \to L^2(SE(2, N))$ be an injective lift operator (not necessarily satisfying (2)). Assume that there exists a residual set $\mathcal{R} \subset L^2(\mathbb{R}^2)$ such that $\tilde{\mathcal{L}}(\mathcal{R}) \cap \mathcal{G}$ is residual. Then, the generalized power-spectrum and bispectrum are weakly complete on $L^2(\mathbb{R}^2)$. Namely, for any $f, g \in \mathcal{R}$ it holds that $\mathrm{BS}_{\tilde{\mathcal{L}}f} = \mathrm{BS}_{\tilde{\mathcal{L}}g}$ if and only if $f = \pi(x, k)g$ for some $(x, k) \in SE(2, N)$.*

*Remark 3* In [38], the authors applied their version of Theorem 1 to a non-left-invariant lift $\tilde{\mathcal{L}}$, called *cyclic lift*. Indeed, for this cyclic lift, when $N$ is odd, it is possible to prove that, for any compact $K \subset \mathbb{R}^2$, there exists a residual set $\mathcal{R} \subset L^2(K)$ satisfying the assumptions of Corollary 1.

Unfortunately, Corollary 1 can never be applied to lifts of the form (2). In fact, as proved in Appendix 3, letting $\omega_f(\lambda) := (\hat{f}(R_{-k}\lambda))_{k=0}^{N-1} \in \mathbb{C}^N$, we have that

$$\widehat{\mathcal{L}f}(T^\lambda) = \omega_\Psi(\lambda)^* \otimes \omega_f(\lambda)^*, \quad (12)$$

where for $v, w \in \mathbb{C}^N$, we let $v^* = (\overline{v_k})_k$ and $(v \otimes w)_{k,h} = v_k \overline{w_h}$, so that $(v \otimes w)u = \langle w, u \rangle v$ for all $u \in \mathbb{C}^N$. This immediately implies that rank $\widehat{\mathcal{L}f}(T^\lambda) \leq 1$ and hence that range $\mathcal{L} \cap \mathcal{G} = \emptyset$ whenever $N > 1$.

### 4.2 Rotational Fourier Descriptors

To bypass the difficulty posed by the non-invertibility of the Fourier transform for lifted functions, we are led to consider the following stronger descriptors.

**Definition 2** The *rotational power-spectrum* and *bispectrum* of $\varphi \in L^2(SE(2, N))$ are the collections of matrices, for any $\lambda, \lambda_1, \lambda_2 \in \mathcal{S}$ and $h \in \mathbb{Z}_N$,

$$\mathrm{RPS}\,\varphi(\lambda, h) := \widehat{\varphi}(T^{R_h\lambda}) \circ \widehat{\varphi}(T^\lambda)^*$$
$$\mathrm{RBS}_\varphi(\lambda_1, \lambda_2, h) := \widehat{\varphi}(T^{R_h\lambda_1}) \otimes \widehat{\varphi}(T^{\lambda_2}) \circ \widehat{\varphi}(T^{\lambda_1} \otimes T^{\lambda_2})^*.$$

As already mentioned in the introduction, the rotational descriptors are invariant only under the action of $\mathbb{Z}_N \subset SE(2, N)$ but not under translations. To avoid this problem, let us fix a compact $K \subset \mathbb{R}^2$ and consider the set $\mathcal{A} \subset L^2(\mathbb{R}^2)$ of functions compactly supported in $K$, with non-zero average[5]. Observe that this is an open and dense subset of $L^2(K)$. We can then define the barycenter $c_f \in \mathbb{R}^2$ of $f \in \mathcal{A}$ as

$$c_f = \frac{1}{\mathrm{avg}\, f} \left( \int_{\mathbb{R}^2} x_1 f(x)\, dx, \int_{\mathbb{R}^2} x_2 f(x)\, dx \right),$$

and the centering operator $\Phi : \mathcal{A} \to \mathcal{A}$ as

$$\Phi f(x) := f(x - c_f). \quad (13)$$

---

[5] Recall that the average of $f : \mathbb{R}^2 \to \mathbb{R}$ is avg $f = \int_{\mathbb{R}^2} f(x)\, dx$, which is always well-defined for $L^2(\mathbb{R}^2)$ functions with compact support.

Then, considering the centered lift $\mathcal{L}_c = \mathcal{L} \circ \Phi$, we have that $\mathcal{L}_c f = \mathcal{L}_c g$ if and only if $g$ is a translate of $f$. In particular,

$$\mathcal{L}_c f = \Lambda(0, k) \mathcal{L}_c g$$
$$\iff f = \pi(x, k) g \quad \text{for some } x \in \mathbb{R}^2.$$

Let us consider the following set of functions.

**Definition 3** Let $\mathcal{R} \subset L^2(\mathbb{R}^2)$ be the set of real-valued functions $f$ supported in $K$, such that $\hat{f}(\lambda) \neq 0$ for a.e. $\lambda \in \mathbb{R}^2$ and the family $\Omega_f = \{S^k \omega_f(\lambda)\}_{k=0}^{N-1}$ is a basis for $\mathbb{C}^N$, if $N$ is odd, or, if $N$ is even, for

$$\mathcal{X} = \{v \in \mathbb{C}^N \mid v(h) = \overline{v(h + N/2)} \quad \forall h \in \mathbb{Z}_N\}.$$

The dependence of this definition on the parity of $N$ comes from the well-known fact that $\hat{f}(\lambda) = \overline{\hat{f}(-\lambda)}$. Indeed, for $N$ even, this implies that $S^k \omega_f(\lambda) \in \mathcal{X}$ for any $k \in \mathbb{Z}_N$. As such, there is no hope for the family $\Omega_f$ to generate the whole $\mathbb{C}^N$.

Finally, we have the following theorem.

**Theorem 2** *For any compact $K \subset \mathbb{R}^2$, if the mother wavelet $\Psi \in \mathcal{R}$, the rotational power-spectrum and bispectrum are weakly complete on $L^2(K) \cap \mathcal{A}$. Namely, the set $\mathcal{R}$ is open and dense in $L^2(K)$ and for any $f, g \in \mathcal{R} \cap \mathcal{A}$ it holds that $\text{RPS}_{\mathcal{L}_c f} = \text{RPS}_{\mathcal{L}_c g}$ and $\text{RBS}_{\mathcal{L}_c f} = \text{RBS}_{\mathcal{L}_c g}$ if and only if $f = \pi(x, k) g$ for some $(x, k) \in SE(2, N)$.*

Here, we content ourselves to present only a sketch of the proof of this result for the case $N$ odd. The parity of $N$ does not introduce essential problems, up to exploit the fact that range $\widehat{Lf}(T^\lambda) \subset \mathcal{X}$ for all $f \in \mathcal{R}$ and $\lambda \in \mathcal{S}$ and that the equivalence $A$ of the Induction–Reduction Theorem quotients nicely to an equivalence between $\mathcal{X} \times \mathcal{X}$ and $\bigoplus_{k \in \mathbb{Z}_N} \mathcal{X}$. However, in order to prove the key technical point (14) we need a much finer study of the properties of circulant operators, which is outside the scope of this work and we defer to a forthcoming paper by Prandi and Gauthier.

*Proof* (Sketch in the case $N$ odd) The fact that $\mathcal{R}$ is open and dense in $L^2(K)$ follows from the same arguments in Lemma 1.

Let $\text{Circ } v$ be the circulant matrix associated with $v$, that is, $\text{Circ } v = [v, Sv, \ldots, S^{N-1}v]$. Then the condition on $\Omega_f$ for $f \in \mathcal{R}$ is equivalent to the invertibility of $\text{Circ } \omega_f(\lambda)$ for an open and dense set of $\lambda$s. By the properties of the Fourier transform on $\mathbb{R}^2$ w.r.t. translations it follows that

$$\omega_{\Phi f}(\lambda) = \text{diag}_k \left( e^{-i \langle \lambda, R_k c_f \rangle} \right) \omega_f(\lambda).$$

This entails that $\text{Circ } \omega_f(\lambda)$ is invertible if and only if $\text{Circ } \omega_{\Phi f}(\lambda)$ is. Hence, the statement is equivalent to the fact

that for any couple $f, g \in \mathcal{R}$ we have $\text{RBS}_{\mathcal{L} f} = \text{RBS}_{\mathcal{L} g}$ if and only if $f = R_k g$ for some $k \in \mathbb{Z}_N$.

The proof is similar to the one of Theorem 1, but with additional technical difficulties. Let $I$ be the set where $\text{Circ } \omega_f(\lambda)$ and $\text{Circ } \omega_g(\lambda)$ are invertible. By assumption $I$ is open and dense. To overcome the non-invertibility of $\widehat{\mathcal{L} f}$ in the definition the candidate intertwining representation $U$, we exploit the invertibility of the circulant matrices $\text{Circ } \omega_f(\lambda)$ and $\text{Circ } \omega_g(\lambda)$ on an open and dense set. Namely, for any $\lambda \in I$ we let

$$U(T^\lambda)^* := \text{Circ } \omega_g(\lambda) \left( \text{Circ } \omega_f(\lambda) \right)^{-1}.$$

By definition, $U(T^\lambda)$ is circulant and $U(T^\lambda)^* S^k \omega_f(\lambda) = S^k \omega_g(\lambda)$ for any $k \in \mathbb{Z}_N$. Moreover, by (12), this is equivalent to

$$\widehat{\mathcal{L} f}(T^{R_k \lambda}) U(T^\lambda) = \widehat{\mathcal{L} g}(T^{R_k \lambda}), \qquad \forall k \in \mathbb{Z}_N.$$

In particular, $\lambda \mapsto U(T^\lambda)$ is constant on orbits $\{R_k \lambda\}_{k \in \mathbb{Z}_N}$. Finally, $U(T^\lambda)$ is unitary as a consequence, e.g., of Theorem 3.

The main difficulty in the proof is now to derive the equivalent of identity (11), that is, that for an open and dense set of couples $(\lambda_1, \lambda_2)$ we have

$$\widehat{\mathcal{L} f}(T^{R_k \lambda_1} \otimes T^{R_k \lambda_2}) U(T^{\lambda_1}) \otimes U(T^{\lambda_2})$$
$$= \widehat{\mathcal{L} g}(T^{R_k \lambda_1} \otimes T^{R_k \lambda_2}), \quad \forall k \in \mathbb{Z}_N. \tag{14}$$

As already mentioned, the proof of this identity requires a deep use of properties of circulant operators, which is outside the scope of this paper. We thus defer it to a forthcoming paper.

Once (14) is known, the statement follows applying the same arguments as those in Theorem 1. Namely,

1. The function $\lambda \mapsto U(T^\lambda)$ is continuous on $I$. This can be done via the same arguments as in Lemma 2.
2. The function $\lambda \mapsto U(T^\lambda)$ can be extended to a continuous function on $\mathcal{S}$ still satisfying (14). This can be done exactly as in Lemma 3.
3. There exists $k \in \mathbb{Z}_N$ such that $U(T^\lambda) = T^\lambda(0, k)$. This is proved following Lemma 4. Indeed, the fact that now $\lambda \mapsto U(T^\lambda)$ is constant on the orbits $\{R_k \lambda\}_{k \in \mathbb{Z}_N}$ implies that the $\varphi_k$s obtained there have to be independent of $k$. Since $\varphi_k(\lambda) = e^{i \langle R_k x_0, \lambda \rangle}$ for some $x_0 \in \mathbb{R}^2$, this implies that $x_0 = 0$ and hence $\varphi_k \equiv 0$. Obviously, this proves that $U(T^\lambda) = S^k = T^\lambda(0, k)$, for some $k \in \mathbb{Z}_N$. □

### 4.3 Practical Computation of the Fourier Descriptors

Here, we present some explicit formulae for the computation of the Fourier descriptors presented in this section.

In the following, we show that, under some assumptions on the mother wavelet $\Psi$, the concrete computation of the generalized power-spectrum and bispectrum and of their rotational counterparts, depend only on the 2D Fourier transform of $f$.

**Theorem 3** *Assume that the mother wavelet $\Psi \in \mathcal{R}$. Then:*

– *For any $f \in \mathcal{R}$, the generalized power-spectrum and bispectrum of $\mathcal{L}f$ are, respectively, determined by the quantities, for a.e. $\lambda, \lambda_1, \lambda_2 \in \mathcal{S}$,*

$$I_1^\lambda(f) = \|\omega_f(\lambda)\|^2 = \sum_{k=0}^{N-1} |\hat{f}(R_{-k}\lambda)|^2$$

$$I_1^{\lambda_1,\lambda_2}(f) = \langle \omega_f(\lambda_1) \odot \omega_f(\lambda_2), \omega_f(\lambda_1 + \lambda_2) \rangle$$
$$= \sum_{k=0}^{N-1} \hat{f}(R_{-k}\lambda_1)\hat{f}(R_{-k}\lambda_2)\overline{\hat{f}(R_{-k}(\lambda_1 + \lambda_2))}.$$

– *For any $f \in \mathcal{A} \cap \mathcal{R}$, the rotational power-spectrum and bispectrum of $\mathcal{L}_c f$ are, respectively, determined by the quantities, for a.e. $\lambda, \lambda_1, \lambda_2 \in \mathcal{S}$ and $h \in \mathbb{Z}_N$,*

$$I_2^{\lambda,h}(f) = \langle \omega_{\Phi f}(R_h \lambda), \omega_{\Phi f}(\lambda) \rangle$$
$$= \sum_{k=0}^{N-1} \overline{\hat{f}(R_{-k+h}\lambda)}\,\hat{f}(R_{-k}\lambda),$$

$$I_2^{\lambda_1,\lambda_2,h}(f) = \langle \omega_{\Phi f}(R_h \lambda_1) \odot \omega_{\Phi f}(\lambda_2), \omega_{\Phi f}(\lambda_1 + \lambda_2) \rangle$$
$$= \sum_{k=0}^{N-1} \hat{f}(R_{-k+h}\lambda_1)\hat{f}(R_{-k}\lambda_2)\overline{\hat{f}(R_{-k}(\lambda_1 + \lambda_2))}.$$

*Here, $\Phi : \mathcal{A} \to \mathcal{A}$ is the centering operator defined in* (13).

*Remark 4* Theorem 3 shows in particular that the result of Theorem 2 is indeed stronger than the completeness result for the generalized bispectrum of the cyclic lift obtained in [38]. Indeed, in that work is proved that the latter (for odd $N$) is determined exactly by the quantities, for a.e. $\lambda_1, \lambda_2 \in \mathcal{S}$ and $h, k \in \mathbb{Z}_N$,

$$\tilde{I}_2^{\lambda_1,\lambda_2,k,h} = \langle \omega_{\phi f}(R_h \lambda_1) \odot \omega_{\phi f}(R_k \lambda_2),$$
$$\omega_{\phi f}(\lambda_1 + R_{h+k}\lambda_2) \rangle.$$

In particular, for each $\lambda_1, \lambda_2 \in \mathcal{S}$ one has to compute $N$ times more quantities than those for the rotational bispectrum.

As a corollary of Theorem 3 we show that, in order to compare the power-spectra and bispectra, it is usually enough to compare only the latter.

**Corollary 2** *Let $\Psi \in \mathcal{R}$ and $f, g \in \mathcal{R} \cap \mathcal{A}$. Then, if $\mathcal{L}f$ and $\mathcal{L}g$ have the same generalized (resp. rotational) bispectrum, they have also the same generalized (resp. rotational) power-spectrum.*

*Proof* We only prove the result for the rotational descriptors. In order to prove the one for the generalized descriptors, it will be enough to fix $h = 0$ in the following. By Theorem 3 it is enough to show that whenever $I_2^{\lambda_1,\lambda_2,h}(f) = I_h^{\lambda_1,\lambda_2,h}(g)$ for a.e. $\lambda_1, \lambda_2 \in \mathcal{S}$ and any $h \in \mathbb{Z}_N$, then $I_2^{\lambda,h}(f) = I_1^{\lambda,h}(g)$ for a.e. $\lambda \in \mathcal{S}$ and any $h \in \mathbb{Z}_N$. We start by observing that by the Paley–Wiener Theorem all these quantities are analytic, since $f$ and $g$ are compactly supported. Moreover,

$$\lim_{\lambda_1,\lambda_2 \downarrow 0} I_2^{\lambda_1,\lambda_2,h}(f) = N\,\hat{f}(0)|\hat{f}(0)|^2 = N\,\mathrm{avg}(f)^3,$$

and the same is true for $g$. Thus, $\mathrm{avg}(f) = \mathrm{avg}(g)$. Finally, the result follows observing that

$$\lim_{\lambda_2 \downarrow 0} I_2^{\lambda_1,\lambda_2,h}(f) = \mathrm{avg}(f)\,I_2^{\lambda_1,h}(f).$$

□

# 5 Experimental Results

The goal of this section is to evaluate the performance of the invariant Fourier descriptors defined in the previous section on a large image database for object recognition. In addition to the generalized power-spectrum (PS) and bispectrum (BS) and the rotational power-spectrum (RPS) and bispectrum (RBS), we also consider the combination of the RPS and BS descriptors. Indeed, combining these two descriptors seems to be a good compromise between the theoretical result of completeness given by Theorem 2, which only holds for the RBS, and computational demands, as the results on the COIL-100 database will show.

After showing how to efficiently compute these descriptors and presenting the image data-set, we analyze some experimental results. In order to estimate the features capabilities, we use a support vector machine (SVM) [39] as supervised classification method. The recognition performances of the different descriptors regarding invariance to rotation, discrimination capability and robustness against noise are compared.

## 5.1 Implementation

As proved in Theorem 3, the equality of the Fourier descriptors we introduced does not depend on the choice of the mother wavelet $\Psi$. Accordingly, in our implementation we only computed the quantities introduced in Theorem 3,
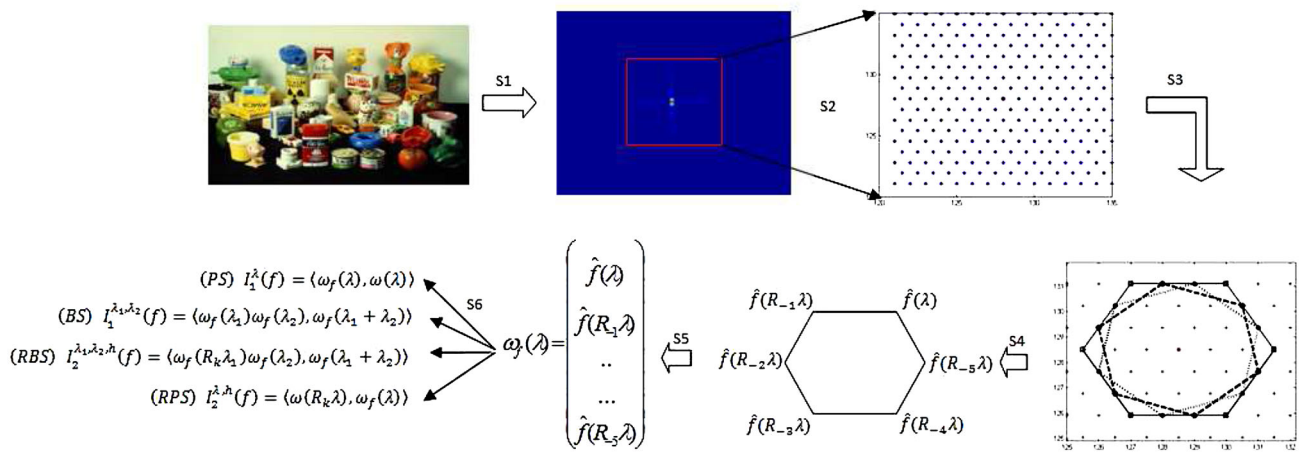
**Fig. 1** Steps of computing the invariant descriptors. (S1) computation of the shifted $FFT$ of the image $f$, (S2) generation of the hexagonal grid, (S3) extraction of different hexagons, (S4) evaluation of the $FFT$ of $f$ on each extracted hexagon, (S5) generation of the vector $\omega_f(\lambda)$ and (S6) computation of the four invariants

whose complexity is reduced to the efficient computation of the vector $\omega_f(\lambda)$, for a given $\lambda \in \mathcal{S}$. We recall that this vector is obtained by evaluating the Fourier transform of $f$ on the orbit of $\lambda$ under the action of discrete rotations $R_{-k}$ for $k \in \mathbb{Z}_N$.

Let us remark that, although in our implementation we chose this approach, in principle fixing a specific mother wavelet could be useful to appropriately weight descriptors depending on the associated frequencies. Indeed, preliminary tests with a Gabor mother wavelet (which can be easily shown to be in $\mathcal{R}$) showed slightly better results at a bigger computational cost.

For the implementation we chose to consider $N = 6$ and to work with images composed of hexagonal pixels. There are two reasons for this choice:

– It is well-known that retinal cells are distributed in a hexagonal grid, and thus it is reasonable to assume that cortical activations reflect this fact.
– Hexagonal grids are invariant under the action of $\mathbb{Z}_6$ and discretized translations, which is the most we can get in the line of the invariance w.r.t. $SE(2, 6)$. Indeed, apart from the hexagonal lattice, the only other lattices on $\mathbb{R}^2$ which are invariant by some $\mathbb{Z}_N$ and appropriate discrete translations are obtained with $N = 2, 3, 4$.

The different steps of computation of the descriptors[6] are described in Fig. 1 and given as follows:

---

[6] MATLAB sample code for the implementation of the rotational bispectral invariants can be found at https://nbviewer.jupyter.org/github/dprn/bispectral-invariant-svm/blob/master/Invariant_computation_matlab.ipynb

**Table 1** Dimension of the feature vectors for the Fourier descriptors under consideration

| Descr. | Dim. |
| --- | --- |
| PS | 136 |
| BS | 717 |
| RPS | 816 |
| RBS | 4417 |
| RPS + BS | 1533 |

1. The input image is converted to grayscale mode, the Fourier transform is computed via FFT, and the zero-frequency component is shifted to the center of the spectrum (Fig. 1, S1).
2. For cost computational reasons and since we are dealing with natural images, for which the relevant frequencies are the low ones, we extract a grid of $16 \times 16$ pixels around the origin (Fig. 1, S2).
3. The invariants of Theorem 3 are computed from the shifted Fourier transform values, on all frequencies in an hexagonal grid inside this $16 \times 16$ pixels square. A bilinear interpolation is applied to obtain the correct values of $\omega_f(\lambda)$ (Fig. 1, S3, S4, S5, S6). The final dimension of the feature vector is given in Table 1.

### 5.2 Test Protocol

We use the Fourier descriptors to feed an SVM classifier, via the MATLAB Statistics and Machine Learning Toolbox, applying it on a database of 7200 objects extracted from the Columbia Object Image Library (COIL-100) and a database of 400 faces extracted from ORL face database. Finally, we

compare the results obtained with those obtained using traditional descriptors.

The result of the training step consists of the set of support vectors determined by the SVM-based method. During the decision step, the classifier computes the Fourier descriptors and the model determined during the training step is used to perform the SVM decision. The output is the image class.

For COIL-100 database, two cases are studied: a case without noise and another with noise. In the first one, tests have been performed using 75% of the COIL-100 database images for training and 25% for testing.In the second one, we have used a learning data-set composed of all the 7200 images (100 objects with 72 views) without noise and a testing dataset composed of 15 randomly selected views per object to which an additive Gaussian noise with $S_d$ of 5, 10 and 20 was added (see Fig. 4).

We evaluate separately the recognition rate obtained using the four previous invariant descriptors and the combination of the RPS & BS invariants to test their complementarity. Then, we compare their performance with the Hu's moments (HM), the Zernike's moments (ZM), the Fourier–Mellin transform (FM), described in Appendix 1, and the local SIFT and HOG descriptors [12] whose performance under the same conditions has been tested in [7],

Since we use the RBF kernel in the SVM classification process, this depends on the kernel size $\sigma$. The results presented here are obtained by choosing empirically the value $\sigma_{opt}$ that provided maximum recognition rate.

## 5.3 Experiments

The performances of the different invariant descriptors are analyzed with respect to the recognition rate given a learning set. Hence, for a given ratio, the learning and testing sets have been built by splitting randomly all examples. Then, due to randomness of this procedure, multiple trials have been performed with different random draws of the learning and testing set. In the case of an added noise, since as mentioned before the learning set comprised all images, this procedure is applied only to the testing set.

The parameters of our experiments are the following:

1. The learning set $c_i$ corresponding to the values of an invariant descriptor computed on an image from the database;
2. The classes $\hat{c}_i \in \{1, 100\}$ corresponding to the object class.
3. Algorithm performance: the efficiency is given through a percentage of the well recognized objects composing the testing set.
4. Number of random trials: fixed to 5.

**Fig. 2** Sample objects of COIL-100 database

5. Kernel K: a Gaussian kernel of bandwidth $\sigma$ is chosen

$$K(x, y) = e^{\frac{-\|x-y\|^2}{2\sigma^2}}$$

$x$ and $y$ correspond to the descriptors vectors of objects.

For solving a multi-class problem, the two most popular approaches are the one-against-all (OAA) method and the one-against-one (OAO) method [29]. For our purpose, we chose an OAO SVM, because it is substantially faster to train and seems preferable for problems with a very large number of classes.

### 5.3.1 COIL-100 Databases

The Columbia Object Image Library (COIL-100, Fig. 2) is a database of color images of 100 different objects, where 72 images of each object were taken at pose intervals of 5°.

### 5.3.2 Classification Performance

Table 2 presents results obtained testing our object recognition method with the COIL-100 database. The best results were achieved using the local SIFT descriptor. The RBS comes in the second place and the local HOG features come third. Indeed it has been demonstrated in the literature, these local methods currently give the best results. However, if noise is added on the image, the use of global approach is better than the use of local ones. The main reason is that the key-points detector used in the local method produce in these cases many key-points that are nor relevant for object recognition. This will be shown in the next subsection.

In Fig. 3 we present the recognition rate as a function of the size of the training set. As expected, this is an increasing function and we remark that the RBS and the combination of the RPS and the BS give better results than the other global invariant descriptors.

**Table 2** Recognition rate for each descriptor using the COIL-100 database. The test results for ZM, HM, FM, and SIFT are taken from [7]

| Descriptors | Recognition rates (%) |
| --- | --- |
| RBS | **95.5** |
| BS | 88 |
| PS | 84.3 |
| RPS | 89.8 |
| RPS+BS | **92.8** |
| ZM | 91.9 |
| HM | 80.2 |
| FM | 89.6 |
| HOG | **95.3** |
| SIFT | **100** |

The best results for each trial are given in bold

### 5.3.3 Robustness Against Noise

Also in this case, test results for ZM, HM, FM, and SIFT are taken from [7].

Results presented in Table 3 show that noise has little influence on classification performance when we use a global descriptor such as RBS, BS, the combination of BS & RPS,

ZM, HM and FM. It has however a sensible effect on the SIFT local descriptor, and a big one on the HOG local descriptor.

### 5.3.4 The ORL Database

The Cambridge University ORL face database (Fig. 5) is composed of 400 gray level images of ten different patterns for each of 40 persons. The variations of the images are across time, size, pose and facial expression (open/closed eyes, smiling/not smiling), and facial details (glasses/no glasses).

In the literature, the protocol used for training and testing is different from one paper to another. In [36], a hidden Markov model (HMM) based approach is used, and the best model resulted in recognition rate of 95%, with high computational cost. In [21], Hjelmas reached a 85% recognition rate using the ORL database and feature vector consisting of Gabor coefficients.

We perform experiments on the ORL database using the RBS, BS, PS, RPS, ZM, HU, FM, and the combination of the RPS & BS descriptors. Since the local descriptors SIFT and HOG obtained, predictably, almost perfect scores, we do not present them. The results are shown in Table 4, where we clearly see that the RBS invariant descriptor gives the best recognition rate $c = 89.8\%$, faring far better than before w.r.t. the combination of RPS and BS descriptors.

**Fig. 3** Classification rate for different size of the training database. The test results for ZM, HM, FM, and SIFT are taken from [7]
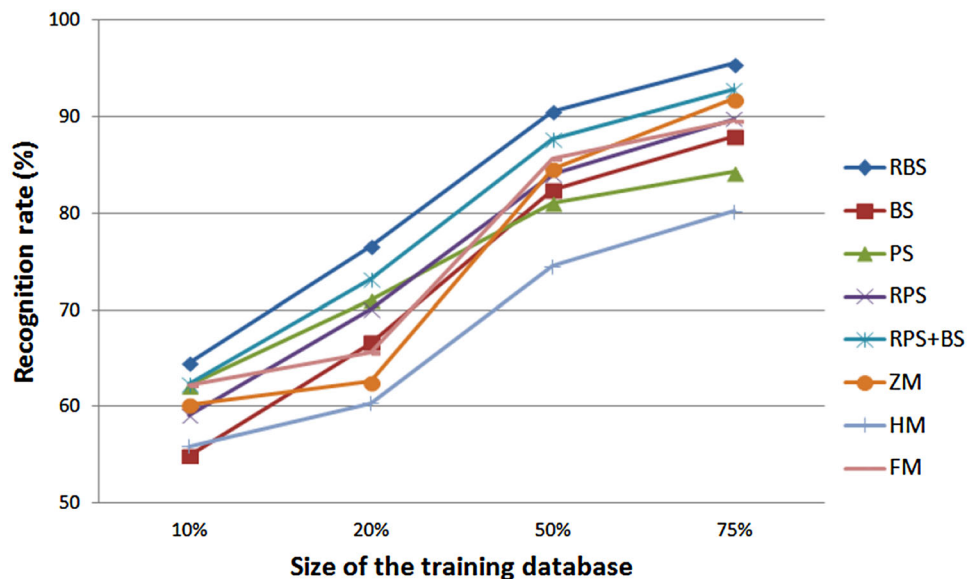


**Table 3** Classification rate on COIL-100 noisy database. The test results for ZM, HM, FM, and SIFT are taken from [7]

| $S_d$ | RBS (%) | BS (%) | PS (%) | RPS (%) | RPS+BS (%) | ZM (%) | HM (%) | FM (%) | SIFT (%) | HOG (%) |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 5 | **100** | **100** | 71.5 | 99.8 | **100** | **100** | 95.2 | 98.6 | 89.27 | 4 |
| 10 | **100** | **100** | 71.2 | 99.8 | **100** | **100** | 95.2 | 95.2 | 88.89 | 1.2 |
| 20 | **100** | **100** | 67.8 | 99.8 | **100** | **100** | 91.4 | 90.2 | 85.46 | 1 |

The best results for each trial are given in bold

**Fig. 4** Sample of COIL-100 noisy object



**Fig. 5** Face samples from the ORL database

**Table 4** Recognition rate for each descriptor using the ORL database

| Descriptors | Recognition rates (%) |
| --- | --- |
| RBS | **89.8** |
| BS | 67.9 |
| PS | 49.2 |
| RPS | 76.9 |
| RPS+BS | 79.8 |
| ZM | 75 |
| HM | 43.5 |
| FM | 47.6 |

The best result for each trial is given in bold

## 6 Conclusion and Perspectives

In this paper, we presented four Fourier descriptors over the semidiscrete roto-translation group $SE(2, N)$. Then, we proved that the generalized power-spectrum (PS) and bispectrum (BS)—and thus the rotational power-spectrum (RPS) and bispectrum (RBS)—are weakly complete, in the sense that they allow to distinguish over an open and dense set of compactly supported functions $\varphi \in L^2(SE(2, N))$ up to the $SE(2, N)$ action. This generalizes a result of [38]. We then considered a framework for the application of these Fourier descriptors to roto-translation invariant object recognition, inspired by some neurophysiological facts on the human primary visual cortex. In this framework, we showed that the rotational bispectrum is indeed a weakly complete roto-translation invariant for planar images. Moreover, although the proposed Fourier descriptors are given in terms of com-

plex mathematical objects, we showed that they can be implemented in a straightforward way as linear combinations of the values of the 2D Fourier transform of the image.

In the second part of the paper, we proposed an evaluation of the performances of these Fourier descriptors in object recognition and we presented the results obtained on different databases: the COIL-100 database, composed of several objects undergoing 3D rotation and scales changes, and the ORL-database, on which different human faces are subject to several kinds of variations. For both these databases, the global Fourier descriptors introduced in this paper are the most efficient global descriptors tested, equalled only, for noisy images, by the Zernike Moments. Although for unperturbed images the local SIFT descriptor gives better recognition rate, the addition of noise leads to the global descriptors outperforming the local ones. These results thus show the rotational bispectrum (RBS) to be a very good Fourier descriptor for object recognition, consistently with the theoretical weak completeness result. When the dimension of the feature vector is an issue, the RBS can be replaced by a combination of the generalized bispectrum (BS) and the rotational power-spectrum (RPS), which yields slightly worse results with a feature vector of length almost one-third.

An extension of the object recognition method presented in this paper to an AdaBoost framework for the problem of object detection is currently ongoing.

## Appendix 1: Moment Invariants and Fourier–Mellin Transform

In this section, following [7], we review the two most used classes of moment invariants, Hu and Zernike, and Fourier–Mellin descriptors, that we use as a comparison for our generalized Fourier descriptors.

Moment invariants were first introduced to the pattern recognition and image processing community in 1962 by Hu [22], with the introduction of the seven Hu moments which are invariants under translation, rotation and scaling. These are derived from a scaling and translation invariant modification of the standard moments of an image $I : \mathbb{R}^2 \to \mathbb{R}$. Namely,

$$v_{p,q} = \frac{u_{p,q}}{u_{0,0}^{(1+\frac{p+q}{2})}},$$

where

$$u_{p,q} = \int_{\mathbb{R}^2} (x - x_0)^p (y - y_0)^q I(x, y) dx dy,$$

and $x_0 = \frac{m_{1,0}}{m_{0,0}}$ and $y_0 = \frac{m_{0,1}}{m_{0,0}}$ are the coordinates of the barycenter computed via the standard $(p+q)$th order moments of $I$:

$$m_{p,q} = \int_{\mathbb{R}^2} x^p y^q I(x,y) dx dy.$$

Another important class of moments are the Zernike ones, introduced in [9] and computed via orthogonal Zernike polynomials. The Zernike moment of order $(m,n)$ is:

$$Z_{mn} = \frac{m+1}{n} \sum_x \sum_y I(x,y) \left[ V_{mn}(x,y) \right],$$

where $x^2 + y^2 < 1$ and $V_{mn}(x,y)$ are the Zernike polynomials defined in polar coordinates as $V_{mn}(r,\theta) = R_{mn}(r)e^{jn\theta}$, where

$$R_{mn}(r) = \sum_{s=0}^{\frac{m-|n|}{2}} \frac{(-1)^s (m-s)! r^{m-2s}}{s! \left( \frac{m+|n|}{2} - s \right)! \left( \frac{m-|n|}{2} - s \right)!}$$

These moments present several advantages. Indeed, beside a rotation and translation invariance they have nice orthogonality properties and are considered to be robust against image noise. In particular, the orthogonality property helps in achieving a near zero value of redundancy measure in a set of moments functions [8].

Finally, strictly related to Fourier descriptors are the descriptors obtained via Fourier–Mellin transform (FMT), presented in [13]. The FMT of an image $I$, that we assume to be given in polar coordinates, is defined as:

$$M_I(u,v) = \frac{1}{2\pi} \int_0^{2\pi} \int_0^{\infty} I(r,\theta) r^{-iv} e^{-iu\theta} \frac{dr}{r} d\theta.$$

Following [13], we will indeed compute the analytical Fourier–Mellin transform (AFMT). That is, we replace $I$ in the above definition with its regularized version $I_\sigma(r,\theta) = r^\sigma I(r,\theta)$, where $\sigma > 0$. Finally, each feature $M_{I_\sigma}(u,v)$ is modified in order to compensate for the rotation, translation and size changes of the object.

## Appendix 2: Auxiliary Lemmata for the Proof of Theorem 1

**Lemma 1** *The set $\mathcal{G}$ introduced in Theorem 1 is open and dense in $L^2(K \times \mathbb{Z}_N)$.*

*Proof* We start by showing that $\mathcal{G} \neq \varnothing$. To this aim, it suffices to consider $\varphi$ such that $\varphi(\cdot, k) \equiv 0$ for all $k \in \mathbb{Z}_N \setminus \{0\}$ and $\varphi(\cdot, 0) \neq 0$ such that $\mathrm{supp}\, \mathcal{F}(\varphi(\cdot, 0)) = \mathbb{R}^2$. By (6), we then have $\varphi \in \mathcal{G}$, since

$$\det \hat{\varphi}(T^\lambda) = \prod_{k \in \mathbb{Z}_N} \mathcal{F}(\varphi(\cdot, 0))(R_{-k}\lambda) \neq 0 \quad \forall \lambda \in \mathcal{S}.$$

For any $\varphi \in \mathcal{G}$ and $k \in \mathbb{Z}_N$, the Paley–Wiener Theorem implies that $\mathcal{F}(\varphi(\cdot, k))$ is analytic. In particular, by (6), $\lambda \mapsto \det \hat{\varphi}(T^\lambda)$ is analytic. Thus, $\varphi \in \mathcal{G}$ if and only if $\varphi(T^{\lambda_0})$ is invertible for some $\lambda_0 \in \mathcal{S}$.

We claim that the set $\mathcal{G}$ is dense. Indeed, let $\varphi \notin \mathcal{G}$ and fix some $\eta \in \mathcal{G}$ and $\lambda_0 \in \mathcal{S}$ such that $\hat{\eta}(T^{\lambda_0})$ is invertible. By analyticity of $\varepsilon \mapsto \det(\hat{\varphi}(T^{\lambda_0}) + \varepsilon \hat{\eta}(T^{\lambda_0}))$ follows that $\varphi + \varepsilon\eta \in \mathcal{G}$ for sufficiently small $\varepsilon > 0$, which entails that $\varphi \in \bar{\mathcal{G}}$, proving the claim.

Let us prove that $\mathcal{G}$ is open in $L^2(K \times \mathbb{Z}_N)$. To this aim, fix $\varphi \in \mathcal{G}$ and $\varphi_n \to \varphi$ in $L^2(K \times \mathbb{Z}_N)$. This implies that $\hat{\varphi}_n \to \hat{\varphi}$ in $L^2(\widehat{SE(2,N)})$, and in particular that $\hat{\varphi}_n \to \hat{\varphi}$ in measure. By definition of convergence in measure, this implies that for sufficiently big $n$ it has to hold $\det \hat{\varphi}_n(T^{\lambda_0}) \neq 0$. Hence $\varphi_n \in \mathcal{G}$ for $n$ sufficiently big and $\mathcal{G}$ is open. □

Before diving into the proofs of the other auxiliary lemmata, we make the following observation. Let $\lambda_1, \lambda_2 \in I$ be such that $\lambda_1 + R_k\lambda_2 \in I$ for all $k \in \mathbb{Z}_N$. Applying the Induction–Reduction theorem (9) to (11) yields

$$
\begin{aligned}
& A \circ U(T^{\lambda_1}) \otimes U(T^{\lambda_2}) \circ A^{-1} \\
&= \bigoplus_{k \in \mathbb{Z}_N} \varphi(T^{\lambda_1 + R_k\lambda_2})^{-1} \eta(T^{\lambda_1 + R_k\lambda_2}) \\
&= \bigoplus_{k \in \mathbb{Z}_N} U(T^{\lambda_1 + R_k\lambda_2}).
\end{aligned}
\tag{15}
$$

**Lemma 2** *The function $\lambda \mapsto U(T^\lambda)$ is continuous on $I$.*

*Proof* Fix $\lambda_0 \in I$ and an open set $V \subset \mathbb{R}^2$ such that

$$\int_V U(T^{\lambda_2})^*_{i,j} d\lambda_2 > 0 \quad \forall i,j \in \mathbb{Z}_N.$$

This is possible since $U \not\equiv 0$. Since the set $I$ is open dense, up to reducing $V$ we can assume that there exists a neighborhood $W$ of $\lambda_0$ such that $V + \lambda \subset I$ for any $\lambda \in W$. Then, (15) holds for $\lambda_1 \in W$ and $\lambda_2 \in V$. Explicitly computing the $0,0$ block of (15), we have

$$U(T^{\lambda_1})_{i,j} U(T^{\lambda_2})^*_{i,j} = U(T^{\lambda_1 + \lambda_2})_{i,j} \quad \forall i,j \in \mathbb{Z}_N.$$

Then, integrating it over $V$ w.r.t. $\lambda_2$ yields

$$U(T^{\lambda_1})_{i,j} = \frac{\int_{V+\lambda} U(T^{\lambda_2})_{i,j} d\lambda_2}{\int_V U(T^{\lambda_2})^*_{i,j} d\lambda_2} \quad \forall \lambda_1 \in W, \forall i,j \in \mathbb{Z}_N$$

Since the function on the r.h.s. is clearly continuous on $W$ this proves the continuity at $\lambda_0$ of $\lambda \mapsto U(T^\lambda)$, completing the proof. □

**Lemma 3** *The function $\lambda \mapsto U(T^\lambda)$ can be extended to a continuous function on $\mathbb{R}^2$ for which (11) is still true.*

*Proof* Let $\lambda_0 \notin I$. Since $I$ is an open and dense set, this implies that $\lambda_0$ is in its closure and that we can choose $\lambda_1, \lambda_2 \in I$ such that $\lambda_0 = \lambda_1 + R_{k_0}\lambda_2$ for some $k_0 \in \mathbb{Z}_N$ and $\lambda_1 + R_k\lambda_2 \in I$ for any $k \neq k_0$. We then let

$$U(T^{\lambda_0}) := \left(A \circ U(T^{\lambda_1}) \otimes U(T^{\lambda_2}) \circ A^*\right)^{k_0,k_0}. \tag{16}$$

We now prove that the above definition does not depend on the choice of $\lambda_1, \lambda_2$ and $k_0$. By openness of $I$, there exists a neighborhood $V$ of $\lambda_2$ entirely contained in $I$. Then, up to taking a smaller $V$, it holds that $\lambda_1 + R_{k_0}\lambda_2' \in I$ for any $\lambda_2' \in V \setminus \{\lambda_2\}$. By (15), this implies that for any $\mu_1 + R_\ell\mu_2 = \lambda_0$ it holds

$$(A \circ U(T^{\lambda_1}) \otimes U(T^{\lambda_2'}) \circ A^*)^{k_0,k_0} =$$
$$(A \circ U(T^{\mu_1}) \otimes U(T^{\mu_2'}) \circ A^*)^{\ell,\ell}.$$

for $\lambda_2'$ and $\mu_2'$ sufficiently near, but different, to $\lambda_2$ and $\mu_2$, respectively. By the continuity of $U$ on $I$, proved in Lemma 2, this implies that this equation has to hold also for $\lambda_2' = \lambda_2$ and $\mu_2' = \mu_2$. Hence, (16) does not depend on the choice of $\lambda_1, \lambda_2$ and $k_0$.

Finally, the fact that $\hat{f}(T^{\lambda_1} \otimes T^{\lambda_2}) \circ U(T^{\lambda_1}) \otimes U(T^{\lambda_2}) = \hat{g}(T^{\lambda_1} \otimes T^{\lambda_2})$ for any $\lambda_1, \lambda_2$ follows from (16) and (15). $\square$

**Lemma 4** *There exists $a \in SE(2, N)$ such that $U(T^\lambda) = T^\lambda(a)$.*

*Proof* By definition of $U$ it holds that

$$\bigoplus_{k \in \mathbb{Z}_N} U(T^{\lambda_1 + R_k\lambda_2}) \circ A = A \circ U(T^{\lambda_1}) \otimes U(T^{\lambda_2}) \; \forall \lambda_1, \lambda_2 \neq 0.$$

Then, for any $i, j, \ell, k$,

$$U(T^{\lambda_1})_{\ell,i} U(T^{\lambda_2})_{\ell-k,j} = \begin{cases} U(T^{\lambda_1 + R_k\lambda_2})_{\ell,i} & \text{if } j = i - k, \\ 0 & \text{otherwise.} \end{cases} \tag{17}$$

By invertibility of $U(T^{\lambda_1})$, there exists $i_0 \in \mathbb{Z}_N$ such that $U(T^{\lambda_1})_{0,i_0} \neq 0$. Using (17) this implies that $U(T^{\lambda_2})_{-k,j} = 0$ for any $j \neq i_0 - k$. Namely, we have proved that there exists a family of functions $\varphi_{-k} : \mathcal{S} \to \mathbb{C}$ such that $U(T^{\lambda_1})_{-k,\cdot} = \varphi_{-k}(\lambda_1) \delta_{i_0-k}$ or, equivalently, that

$$U(T^\lambda) = \text{diag}_k \varphi_k(\lambda) S^{i_0}.$$

By the explicit expression (4) of $T^\lambda$, in order to complete the proof it suffices to prove that $\varphi_k(\lambda) = e^{i\langle x_0, R_{-k}\lambda\rangle}$ for some $x_0 \in \mathbb{R}^2$.

By continuity and unitarity of $U$, the $\varphi_k$s are continuous and satisfy $|\varphi_h(\lambda)| = 1$. Using again (17) with $j = i_0 - k$, we obtain

$$\varphi_\ell(\lambda_1 + R_k\lambda_2) = \varphi_\ell(\lambda_1)\varphi_{\ell-k}(\lambda_2), \tag{18}$$

for any $\lambda_1, \lambda_2 \neq 0$ and $\ell, k \in \mathbb{Z}_N$.

We claim that the $\varphi_\ell$'s are characters of $\mathbb{R}^2$. Indeed, let us fix $k = 0$ in (18):

$$\varphi_\ell(\lambda_1 + \lambda_2) = \varphi_\ell(\lambda_1)\varphi_\ell(\lambda_2). \tag{19}$$

Choosing $\lambda_2 = -\lambda_1$ in the above shows that $\varphi_\ell$ can be extended at 0. Moreover, letting $\lambda_1 = 0$ and taking the limit $\lambda_2 \to 0$ shows that this extension is continuous. Since characters of $\mathbb{R}^2$ are exactly the continuous functions satisfying (19), the claim is proved.

By Pontryiagin duality, there exists $x_\ell \in \mathbb{R}^2$ such that $\varphi_\ell(\lambda) = e^{i\langle\lambda, x_\ell\rangle}$. Finally, by (18) with $k \in \mathbb{Z}_N$ one obtains that $R_{-k}x_\ell = x_{\ell-k}$, which proves that there exists $x_0 \in \mathbb{R}^2$ such that $\varphi_\ell(\lambda) = e^{i\langle x_0, R_{-k}\lambda\rangle}$. This completes the proof of the statement. $\square$

## Appendix 3: Proofs

*Proof* (Formula (12)) Let $\lambda \in \mathcal{S}$ and consider $v \in \mathbb{C}^N$. Observe that $(x, k)^{-1} = (-R_{-k}x, -k)$. Then, by (5), (2), and (4), for any $h \in \mathbb{Z}_N$ we have

$$(\widehat{\mathcal{L}}(T^\lambda).v)_h = \sum_{k=0}^{N-1} \int_{\mathbb{R}^2} \mathcal{L}(x, k) e^{-i\langle\lambda, R_{h-k}x\rangle} v_{h-k} \, dx$$

$$= \sum_{k=0}^{N-1} v_{h-k} \int_{\mathbb{R}^2}\int_{\mathbb{R}^2} f(y)\bar{\Psi}(R_{-k}(y-x)) e^{-i\langle\lambda, R_{h-k}x\rangle} \, dy \, dx$$

$$= \sum_{k=0}^{N-1} v_{h-k} \int_{\mathbb{R}^2}\int_{\mathbb{R}^2} \bar{\Psi}(z) f(y) e^{-i\langle R_{k-h}\lambda, y-z\rangle} \, dy \, dz$$

$$= \bar{\hat{\Psi}}(R_h\lambda) \sum_{k=0}^{N-1} v_{h-k} \hat{f}(R_{k-h}\lambda)$$

$$= \overline{\omega_\Psi(\lambda)}_h \langle\omega_f(\lambda), v\rangle.$$

By definition of $\omega_\Psi(\lambda)^* \otimes \omega_f(\lambda)^*$; this completes the proof. $\square$

In order to prove Theorem 3, we need the following explicit description of the equivalence in the Induction–Reduction theorem of Sect. 3.1.

**Lemma 5** *For any $M, N \in \mathbb{C}^{N \times N}$ we have*

$$\left(A \circ (M \otimes N) \circ A^{-1}\right)^{k,h} = (M_{i,j} N_{i-k,j-h})_{i,j \in \mathbb{Z}_N}.$$

*Proof* Observe that for any $\mathbf{v} \in \mathbb{C}^N \otimes \mathbb{C}^N$ it holds $(M \otimes N).\mathbf{v} = M \circ \mathbf{v} \circ N^T$. Thus,

$$[A \circ (M \otimes N).\mathbf{v}]_k(i) = \sum_{j,\ell=0}^{N-1} M_{i,j}\, N_{i-k,\ell}\, \mathbf{v}(j,\ell).$$

Since is straightforward to check that $A^{-1} : \bigoplus_{k \in \mathbb{Z}_N} \mathbb{C}^N \to \mathbb{C}^{N \times N}$ is given by $[A^{-1}(w_\ell)_{\ell \in \mathbb{Z}_N}](k,h) = w_{k-h}(k)$, we then have

$$
\begin{aligned}
&[A \circ (M \otimes N) \circ A^{-1}.(w_h)_{h \in \mathbb{Z}_N}]_k(i) \\
&= \sum_{j,\ell=0}^{N-1} M_{i,j}\, N_{i-k,\ell}\, w_{j-\ell}(j) \\
&= \sum_{j,h=0}^{N-1} M_{i,j}\, N_{i-k,j-h}\, w_h(j).
\end{aligned}
$$

By (8), the proof is completed. $\qquad\square$

*Proof* (Proof of Theorem 3) Without loss of generality we can restrict ourselves to consider functions such that $\phi f = f$ and $\mathcal{L}_c f = \mathcal{L} f$. We start by the trivial remark that the result on the rotational descriptors contains the one on the generalized ones.

Let us consider

$$
\begin{aligned}
I_2^{\lambda,h}(f) &:= \langle \omega_f(R_h \lambda), \omega_f(\lambda) \rangle, \\
I_2^{\lambda_1,\lambda_2,h}(f) &:= \langle \omega_f(R_h \lambda_1) \odot \omega_f(\lambda_2), \omega_f(\lambda_1 + \lambda_2) \rangle.
\end{aligned}
$$

Since $f$ is assumed to be compactly supported, its Fourier transform is analytic, and so are the functions $\lambda \mapsto I_2^{\lambda,h}(f)$ and $(\lambda_1, \lambda_2) \mapsto I_2^{\lambda_1,\lambda_2,h}(f)$ for any $h \in \mathbb{Z}_N$. Thus, the statement of the proposition reduces to show that $\mathrm{RPS}_{\mathcal{L}f} = \mathrm{RPS}_{\mathcal{L}g}$ (resp. $RB_{\mathcal{L}f} = RB_{\mathcal{L}g}$) if and only if $I_2^{\lambda,h}(f) = I_2^{\lambda,h}(g)$ (resp. $I_2^{\lambda_1,\lambda_2,h}(f) = I_2^{\lambda_1,\lambda_2,h}(g)$) for a.e. $\lambda, \lambda_1, \lambda_2 \in \mathcal{S}$ and all $h \in \mathbb{Z}_N$

Let us recall the following properties of the tensor product, valid for all $v, v_1, v_2, w, w_1, w_2 \in \mathbb{C}^N$:

1. $(v \otimes w)^* = v \otimes w$,
2. $(v_1 \otimes w_1) \circ (v_2 \otimes w_2) = \langle w_1, v_2 \rangle\, v_1 \otimes w_2$,

By these and (12), we immediately have

$$
\begin{aligned}
\mathrm{RPS}_{\mathcal{L}_c f}(\lambda, h) &= \langle \omega_{\phi f}(R_h \lambda)^*, \omega_{\phi f}(\lambda)^* \rangle\, \omega_\Psi(R_h \lambda)^* \otimes \omega_\Psi(\lambda)^* \\
&= I_2^{\lambda,h}(f)\, \omega_\Psi(R_h \lambda)^* \otimes \omega_\Psi(\lambda)^*.
\end{aligned}
$$

Hence, whenever $\omega_\Psi(R_h \lambda)^* \otimes \omega_\Psi(\lambda)^* \neq 0$, $\mathrm{RPS}_{\mathcal{L}_c f}(\lambda, h) = \mathrm{RPS}_{\mathcal{L}_c g}(\lambda, h)$ if and only if $I_2^{\lambda,h}(f) = I_2^{\lambda,h}(g)$. Since $\omega_\Psi(R_h \lambda)^* \otimes \omega_\Psi(\lambda)^* \neq 0$ if and only if $\omega_\Psi(\lambda) \neq 0$, by the fact that $\Psi \in \mathcal{R}$ this is true for a.e. $\lambda \in \mathcal{S}$. This completes the proof of the part of the statement regarding the rotational power-spectrum.

To prove the statement regarding the rotational bispectrum, let $\mathcal{B}_f = A \circ \mathrm{RBS}_{\mathcal{L}f}(\lambda_1, \lambda_2, h) \circ A^{-1}$, where $A$ is the equivalence given by the Induction–Reduction Theorem and defined in (10). Since $A$ is invertible, determining $\mathrm{RBS}_{\mathcal{L}f}(\lambda_1, \lambda_2, h)$ is equivalent to determining $\mathcal{B}_f$. Exploiting the fact that the r.h.s. of (9) is a diagonal matrix, we have

$$
\begin{aligned}
\mathcal{B}_f^{k,\ell} &= \left( A \circ \widehat{\mathcal{L}f}(T^{R_h \lambda_1}) \otimes \widehat{\mathcal{L}f}(T^{\lambda_1}) \circ A^{-1} \right)^{k,\ell} \\
&\qquad\qquad \circ \widehat{\mathcal{L}f}(T^{\lambda_1 + R_\ell \lambda_2})^*.
\end{aligned}
$$

By Lemma 5, formula (12), and explicit computations, we then get

$$
\begin{aligned}
\mathcal{B}_f^{k,\ell} &= \langle \omega_f(R_h \lambda_1) \odot \omega_f(R_\ell \lambda_2), \omega_f(\lambda_1 + R_\ell \lambda_2) \rangle \\
&\quad \times \left( \omega_\Psi(R_h \lambda_1)^* \odot \omega_\Psi(R_{h+k} \lambda_2)^* \right) \otimes \omega_\Psi(\lambda_1 + R_\ell \lambda_2).
\end{aligned}
$$

Similarly to before, $(\omega_\Psi(R_h \lambda_1)^* \odot \omega_\Psi(R_{h+k} \lambda_2)^*) \otimes \omega_\Psi(\lambda_1 + R_\ell \lambda_2) \neq 0$ for a.e. $\lambda_1, \lambda_2 \in \mathcal{S}$ since $\Psi \in \mathcal{R}$. For these couples, $\mathcal{B}_f = \mathcal{B}_g$ if and only if

$$
\begin{aligned}
&\langle \omega_f(R_h \lambda_1) \odot \omega_f(R_\ell \lambda_2), \omega_f(\lambda_1 + R_\ell \lambda_2) \rangle \\
&= \langle \omega_g(R_h \lambda_1) \odot \omega_g(R_\ell \lambda_2), \omega_g(\lambda_1 + R_\ell \lambda_2) \rangle.
\end{aligned}
$$

Finally, making the change of variables $R_\ell \lambda_2 \mapsto \lambda_2$ completes the proof of the theorem. $\qquad\square$

## References

1. Barut, A., Rączka, R.: Theory of Group Representations and Applications. World Scientific, Singapore (1977)
2. Bay, H., Tuytelaars, T., Van Gool, L.: Surf: speeded up robust features. Computer Vision—ECCV 2006, pp. 404–417. Springer, Heidelberg (2006)
3. Boscain, U., Duplaix, J., Gauthier, J.P., Rossi, F.: Anthropomorphic image reconstruction via hypoelliptic diffusion. SIAM J. Control Optim. **50**, 1–25 (2012)
4. Boscain, U., Chertovskih, R., Gauthier, J.P., Remizov, A.: Hypoelliptic diffusion and human vision: a semi-discrete new twist on the petitot theory. SIAM J. Imaging Sci. **7**(2), 669–695 (2014a)
5. Boscain, U., Gauthier, J.P., Prandi, D., Remizov, A.: Image reconstruction via non-isotropic diffusion in Dubins/Reed–Shepp-like control systems. In: 53rd IEEE Conference on Decision and Control, pp. 4278–4283 (2014b)
6. Bressloff, P., Cowan, J., Golubitsky, M., Thomas, P., Wiener, M.: Geometric visual hallucinations, euclidean symmetry and the functional architecture of striate cortex. Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci. **356**, 299–330 (2001)
7. Choksuriwong, A., Emile, B., Rosenberger, C., Lauren, H.: Comparative study of global invariant descriptors for object recognition. J. Electron. Imaging **17**, 1–35 (2008)
8. Chong, C., Raveendran, P., Mukundan, R.: A comparative analysis of algorithms for fast computation of Zernike moments. Pattern Recogn. **36**(3), 731–742 (2003a)
9. Chong, C.W., Raveendran, P., Mukundan, R.: Translation invariants of zernike moments. Pattern Recogn. **36**(8), 1765–1773 (2003b)

10. Chou, J., O'Neill, W., Cheng, H.: Pavement distress classification using neural networks. In: 1994 IEEE International Conference on Systems, Man, and Cybernetics, 1994. Humans, Information and Technology, vol. **1**, pp. 397–401. doi:10.1109/ICSMC.1994.399871 (1994)

11. Citti, G., Sarti, A.: A cortical based model of perceptual completion in the roto-translation space. Pattern Recogn. **24**(3), 307–326 (2006)

12. Dalal, N., Triggs, B.: Histograms of oriented gradients for human detection. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2005. CVPR 2005, IEEE, vol. **1**, pp. 886–893 (2005)

13. Derrode, S., Ghorbel, F.: Robust and efficient fourier-mellin transform approximations for invariant grey-level image description and reconstruction. Comput. Vis. Image Underst. **83**(1), 57–78 (2001)

14. Dubnov, S., Tishby, N., Cohen, D.: Polyspectra as measures of sound texture and timbre. Comput. Vis. Image Underst. **26**(4), 277–314 (1997)

15. Duits, R., Franken, E.: Left-invariant parabolic evolutions on SE(2) and contour enhancement via invertible orientation scores. Part I: Linear left-invariant diffusion equations on SE(2). Q. Appl. Math. **68**, 255–292 (2010)

16. Duits, R., Franken, E.: Left-invariant parabolic evolutions on SE(2) and contour enhancement via invertible orientation scores. Part II: Nonlinear left-invariant diffusions on invertible orientation scores. Q. Appl. Math. **68**, 1–38 (2010)

17. Führ, H., Mayer, M.: Continuous wavelet transforms from semi-direct products : Cyclic representations and Plancherel measure. J. Fourier Anal. Appl. 8(4):1–23. http://www.springerlink.com/index/G7TC4AANGTUC4HXW.pdf, 0102002v1 (2002)

18. Galerne, B., Gousseau, Y., Morel, J.: Random phase textures: theory and synthesis. IEEE Trans. Image Process. **20**(1), 257–267 (2011)

19. Granlund, G.H.: Fourier preprocessing for hand print character recognition. IEEE Trans. Image Process. **C–21**(2), 195–201 (1972)

20. Hewitt, E., Ross, K.: Abstract Harmonic Analysis—Volume 1. Springer, Berlin/New York (1963)

21. Hjelmas, E., Low, B.: Face detection: a survey. Comput. Vis. Image Underst. **83**(3), 236–274 (2001)

22. Hu, M.: Visual pattern recognition by moment invariants. Comput. Vis. Image Underst. **8**(2), 179–187 (1962)

23. Hubel, D., Wiesel, T.: Receptive fields of single neurones in the cat's striate cortex. Comput. Vis. Image Underst. **148**, 574–591 (1959)

24. Kakarala, R.: The bispectrum as a source of phase-sensitive invariants for fourier descriptors: a group-theoretic approach. Comput. Vis. Image Underst. **44**(3), 341–353 (2012)

25. Ke, Y., Sukthankar, R.: Pca-sift: A more distinctive representation for local image descriptors. In: Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004. CVPR 2004. IEEE, vol **2**, p. II-506 (2004)

26. Kuhl, F.P., Giardina, C.R.: Elliptic Fourier features of a closed contour. Comput. Vis. Image Underst. **18**(3), 236–258 (1982)

27. Lowe, D.: Distinctive image features from scale-invariant keypoints. Comput. Vis. Image Underst. **60**(2), 91–110 (2004)

28. Mikolajczyk, K., Schmid, C.: A performance evaluation of local descriptors. Comput. Vis. Image Underst. **27**(10), 1615–1630 (2005)

29. Milgram, J., Cheriet, M., Sabourin, R.: "One Against One" or "One Against All": which one is better for handwriting recognition with SVMs? In: Lorette, G. (ed.) Tenth International Workshop on Frontiers in Handwriting Recognition, Université de Rennes 1, Suvisoft, La Baule. http://www.suvisoft.com (2006)

30. Morel, J.M., Yu, G.: Asift: a new framework for fully affine invariant image comparison. Comput. Vis. Image Underst. **2**(2), 438–469 (2009)

31. Nene, S.A., Nayar, S.K., Murase, H., et al.: Columbia object image library (coil-20). Tech. rep., technical report CUCS-005-96 (1996)

32. Petitot, J.: Neurogéométrie de la vision - Modèles mathématiques et physiques des architectures fonctionnelles. Les Éditions de l'École Polytechnique (2008)

33. Prandi, D., Boscain, U., Gauthier, J.P.: Image processing in the semidiscrete group of rototranslations. In: To appear in the Proceedings of the 2nd Conference on Geometric Science of Information (2015)

34. Raja, D.M.S., Shanmugam, A.: Artificial neural networks based war scene classification using invariant moments and glcm features: A comparative study. Int. J. Eng. Sci. Technol. **3**(2), 1189–1195 (2011)

35. Rajasekaran, S., Pai, G.V.: Image recognition using simplified fuzzy artmap augmented with a moment based feature extractor. Int. J. Pattern Recognit. Artif. Intell. **14**(08), 1081–1095 (2000)

36. Samaria, F., Harter, A.: Parameterisation of a stochastic model for human face identification. In: Proceedings of the Second IEEE Workshop on Applications of Computer Vision, 1994, pp. 138–142 (1994)

37. Sifre, L., Mallat, S.: Rotation, scaling and deformation invariant scattering for texture discrimination. In: 2013 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 1233–1240. doi:10.1109/CVPR.2013.163 (2013)

38. Smach, F., Lemaître, C., Gauthier, J.P., Miteran, J., Atri, M.: Generalized Fourier descriptors with applications to objects recognition in SVM context. Int. J. Pattern Recognit. Artif. Intell. **30**, 43–71 (2008)

39. Vapnik, V.N., Vapnik, V.: Statistical Learning Theory, vol. 1. Wiley, New York (1998)

40. Zahn, C.T., Roskies, R.Z.: Fourier descriptors for plane closed curves. IEEE Trans. Comput. **C–21**(3), 269–281 (1972)
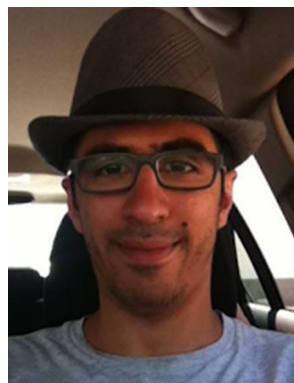
**Amine Bohi** was born in Taza, Morocco, in 1987. He received the Master degree in computer science, computer vision and imaging from the University Sidi Mohammed Ben Abdellah - Faculty of science, Fez, Morocco, in 2011. He currently is a Ph.D. student at the University of South Toulon-Var (France). He belongs to the Laboratory for Systems and Information Science. His research interests concern object detection for maritime surveillance applications.



**Dario Prandi** was born in 1986. He received is Ph.D. in applied mathematics from École Polyechnique, Palaiseau, France, and SISSA, Trieste, Italy, in 2014. He is currently a post-doctoral student at CEREMADE, Univ. Paris-Dauphine. His research interests include sub-Riemannian geometry, image processing and neuroscience.
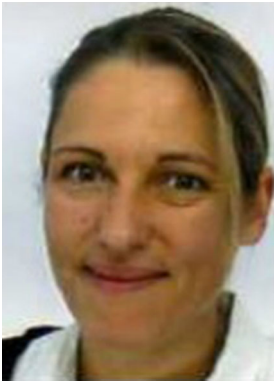
**Vincente Guis** She received her Ph.D. degrees in signal and image processing applied to quality control of mammographic devices and to computer-aided diagnosis for breast cancer from the Aix-Marseille of University with ARCADES, Marseille, France, in 2002. From 2002 to 2009, she worked as a signal and image processing expert for research and development teams of different companies specializing in Astrophysics (Corot Project-CNES-LAM), Geophysics (Magnitude joint venture of CGG Veritas Baker Hughes) and embedding image systems (Shaktiware). In 2010, she joined the LSIS Laboratory (SIIM team) at the University of Toulon, France, as a Research Engineer. Her current fields of interest are Physics, Informatics, Signal and Image Processing, and applied systems and she is an active member of the "Amicale Photon Matière" since 1999.

**Jean-Paul Gauthier** was born in 1952. He is currently a Professor at the University of Toulon (France) in the Dpt of Electrical Engineering. He got his Ph.D. in physics in 1982. He got the medal of "Institut Universitaire de France" in 1992. He was a member of this institute from 1992 to 97, and is now honorary member. He got a Featured review of the American Mathematical Society in 2002, for his work on the subanalyticity of Carnot–Caratheodory distances. His current fields of interest are Automatic Control, Robotics, Signal and Image Processing, and Deterministic Observation Theory (he wrote a reference book at Cambridge University Press in 2001 on this last topic). Besides his academic activities, he is a member of "Federation Française de Go", "Ligue des Libres penseurs" and a strong supporter of "Toulon Rugby Club".

**Frédéric Bouchara** received the DEA (M.Sc degree) from the University of Marseille (France) and the Ph.D. degree from the University of Toulon (France) in 1992 and 1998, respectively. He is currently an associate professor of computer sciences at the University of Toulon. His research focuses on computer vision and image processing, mainly for applications in maritime surveillance and document image processing.