

The Fourier-Argand Representation: An Optimal Basis of Steerable Patterns

Tianle Zhao¹, *Student Member, IEEE*, and Thierry Blu², *Fellow, IEEE*

Abstract—Computing the convolution between a 2D signal and a corresponding filter with variable orientations is a basic problem that arises in various tasks ranging from low level image processing (e.g. ridge/edge detection) to high level computer vision (e.g. pattern recognition). Through decades of research, there still lacks an efficient method for solving this problem. In this paper, we investigate this problem from the perspective of approximation by considering the following problem: what is the optimal basis for approximating all rotated versions of a given bivariate function? Surprisingly, solely minimising the L^2 -approximation-error leads to a rotation-covariant linear expansion, which we name Fourier-Argand representation. This representation presents two major advantages: 1) rotation-covariance of the basis, which implies a “strong steerability” — rotating by an angle α corresponds to multiplying each basis function by a complex scalar $e^{-ik\alpha}$; 2) optimality of the Fourier-Argand basis, which ensures a few number of basis functions suffice to accurately approximate complicated patterns and highly direction-selective filters. We show the relation between the Fourier-Argand representation and the Radon transform, leading to an efficient implementation of the decomposition for digital filters. We also show how to retrieve accurate orientation of local structures/patterns using a fast frequency estimation algorithm.

Index Terms—Fourier-Argand representation, rotation-covariant function, Radon transform, ridge/edge detection, pattern matching.

I. INTRODUCTION

PERFORMING *steerable convolution* — convolving two two-dimensional (2D) signals while allowing one of the signal to be rotated by an arbitrary angle — is a basic requirement for solving many problems in the fields of image processing and computer vision. For instance, one may use derivatives of Gaussian as filters to detect image local structures, e.g., ridges and edges. In his seminal paper [1], Canny used the first derivative of Gaussian to detect edges. Following this idea, many edge/ridge/wedge detection algorithms have been proposed [2]–[5]. By using other directional filters,

Manuscript received August 9, 2019; revised February 10, 2020 and April 17, 2020; accepted April 20, 2020. Date of publication May 6, 2020; date of current version May 29, 2020. This work was supported by the Hong Kong Research Grant Council under Grant #CUHK14210617. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Dong Tian. (*Corresponding author: Tianle Zhao.*)

¹The authors are with the Department of Electronic Engineering, The Chinese University of Hong Kong, Hong Kong (e-mail: tlzhao@ee.cuhk.edu.hk; tblu@ee.cuhk.edu.hk).

²This article has supplementary downloadable material available at HYPERLINK "https://ieeexplore.ieee.org" http://ieeexplore.ieee.org, provided by the authors.

Digital Object Identifier 10.1109/TIP.2020.2990483

computing steerable convolution also serves as a building block for tasks such as image segmentation [6], blood vessel detection [7], texture analysis [8], and rotation-invariant feature extraction [9]. More generally, the problem of pattern matching under in-plane rotations can also be considered as a steerable convolution problem [10].

Looking from the perspective of matched filtering, performing a steerable convolution amounts to solving a *steerable* matched filtering problem [4]. Specifically, for a given 2D signal, e.g., an image $I(x, y)$, and a pattern $h(x, y)$ acting as a filter, the problem is to calculate the convolution of I and rotated versions of h , i.e., steerable convolution consists in calculating

$$(I * {}^\alpha h)(x, y), \quad (1)$$

for all α , where $*$ denotes the 2D convolution, and ${}^\alpha f$ denotes the counter-clockwise rotation of f by the angle α , defined as follows,

$${}^\alpha f(x, y) = f(x \cos \alpha + y \sin \alpha, -x \sin \alpha + y \cos \alpha). \quad (2)$$

It is well known that the matched filter is the optimal linear filter, which maximises the signal-to-noise ratio (SNR) under additive white Gaussian noise [11]. However, calculating the convolution (1) for all α remains difficult, since α takes values continuously on the interval $[0, 2\pi)$, and hence involves uncountably many rotated versions of h .

An intuitive approximate solution is to quantise the angle α , and convolve I with h rotated by the quantised angles. After that, one can interpolate between the quantisation levels. However, this approach requires a large number — inversely proportional to the desired angle accuracy — of filters, which makes this naive solution extremely time consuming and not particularly useful, especially when a relatively high angle accuracy is required.

A. Rotation Invariant Filter Bases

A more elegant solution was proposed by Freeman and Adelson in their pioneering paper [3], where they showed that (1) can be computed efficiently for a class of filters, namely the steerable filters. To be specific, they proved that if a filter h has limited angular frequencies, then the space consisting of all rotated versions of h is of finite dimension. Therefore, any rotated version of h can be represented as a linear combination of a fixed set of basis filters. For such a steerable filter with N distinct angular frequencies, they used N rotated versions of the original filter h as a basis.

As a benefit of this representation, rotating the filter h amounts to changing the coefficients of the linear combination, i.e.,

$${}^a h(x, y) = \sum_{n=0}^{N-1} c_n(\alpha) {}^a h_n(x, y). \quad (3)$$

This representation efficiently solves the problem caused by rotations as long as the filter h satisfies the steerability assumption — having finite number of distinct angular frequencies. In practice, however, the filter h is frequently not limited in angular frequency. Therefore, the authors of [3] proposed to use a bivariate polynomial $P(x, y)$ of degree- K multiplied by an isotropic window $W(\sqrt{x^2 + y^2})$ to approximate h , i.e.,

$$h(x, y) \approx P(x, y)W(\sqrt{x^2 + y^2}). \quad (4)$$

It is not hard to see that this approximation space is rotation-invariant. Indeed, a rotation of the polynomial $P(x, y)$ is still a polynomial of the same degree. Besides, the dimensionality of this approximation space does not exceed $2K + 1$. To see this, one can express the polynomial $P(x, y)$ in polar coordinates, and then use Euler's formula. The steerability of the windowed polynomial approximation (4) ensures that any rotation can be efficiently handled by (3). There have been many follow-ups of this idea. For instance, Jacob *et al.* used derivatives of an isotropic Gaussian function and incorporated Canny-like criterion in [4], where they devised a successful algorithm for the detection of edges, ridges and wedges.

Despite the above advantages, the steerable filter [3] suffers from the following shortcomings. Firstly, the windowed polynomial approximation (4) is not accurate enough, when h is rarely limited in angular frequency. The number of basis filters required to achieve an accurate approximation quickly becomes very large when h becomes less isotropic, which makes the algorithm significantly less efficient. Moreover, using a high-degree polynomial causes numerical stability issues and heavily increases the computational cost. These limitations restrict the steerable filter algorithm from representing highly direction-selective filters, e.g., extremely elongated functions, or general patterns that have complicated structures. Secondly, the linear coefficients $c_n(\alpha)$ in (3) are usually complicated functions of α — high-degree polynomials of trigonometric functions. This makes the inverse problem — retrieving the rotation back from the coefficients — difficult and inaccurate, especially in the presence of noise.

B. The Proposed Approach

We address this steerable convolution problem from the perspective of approximation. By minimising the average L^2 -error when approximating rotated versions of a target pattern/filter, we find the optimal basis consisting of rotation-covariant functions. In addition to the optimality, this representation exhibits other merits that will be discussed in details in Sec. II-C.

Mathematically, the derived representation itself is by no means new. For examples, its polar form is known as the circular harmonic expansion in the optical society [12], which

TABLE I
DISTINCTIVE PROPERTIES OF FOURIER-ARGAND FILTERS

1. Steerability + rotation-covariance ^a ;
2. Minimises the number of filters needed for a given L^2 -error ^b ;
3. On-the-fly calculation for arbitrary patterns ^c .

^a Fourier-Argand filters share steerability with the steerable filter [4], but their rotation-covariance simplifies significantly the relation between the rotation angle and the filter coefficients. (See Sec. II-C1.)

^b Fourier-Argand filters reach the optimal (see Sec. II-C2), while the others need significantly more filters to reach the same L^2 -error. (See Figs. 4 and 5.)

^c See Sec. II-D. We are not aware of any attempt to represent *arbitrary patterns* with other steerable filter bases.

was used to construct rotation-invariant descriptors for pattern recognition [13]. However, only the magnitude of the filter response was used, while the phase was discarded. Despite the wide knowledge that the phase contains important information, a systematic and robust incorporation of the phase information into the descriptor was considered difficult [14]. The importance of the phase is further outlined by the fact that the magnitude of the filter response is very sensitive to different kinds of distortions, such as illumination changes, noise and shifts [15]. Indeed, we show in this paper it is the phase that determines the local orientation of the image (direction of the ridge/edge, pattern orientation), and that maximising over all rotation angles gives a very robust rotation invariant that can be exploited for pattern recognition. The Fourier-Argand representation is also related to other representations that share the same angular part but differ in their radial parts, e.g., the Zernike polynomials used to construct invariants [16] and the basis functions for the polar Fourier transform developed in [14]. The difference is that our representation is optimal for the given pattern.

We would like to clarify the differences between the proposed Fourier-Argand representation in this paper and the Fourier-Argand moments proposed in our previous paper [17]. Firstly, the two are based on different assumptions: the Fourier-Argand moments are only applicable to laminar signals, i.e., 2D signals that only vary in one direction; but, the Fourier-Argand representation is applicable to arbitrary square integrable 2D signals. Secondly, for any given pattern, the Fourier-Argand representation is the optimal linear representation; but, the Fourier-Argand moments (if applicable) are not optimised and do not form a decomposition of the pattern.

In the rest of the paper, we will give the explicit representation of filters that we call Fourier-Argand (Sec. II-A and II-B), and will investigate its properties (Sec. II-C), the most distinctive of which are summarised in Table I. Then we will show how it is possible to compute the Fourier-Argand basis filters using the Radon transformation (Sec. II-D). In Sec. II-E, we will calculate the Fourier-Argand representation of an elongated Gaussian, and demonstrate the significantly higher efficiency of this representation compared to classic steerable filters. Thanks to the particularly simple relation between rotation angle and coefficients of the Fourier-Argand representation, we are able to propose an efficient and accurate algorithm for finding the best

local orientation of patterns in an image (Sec. III): we will exemplify the accuracy of this algorithm on a basic pattern detection problem. We further perform different experiments (blood vessel segmentation, ridge detection and letter matching) over sequences of images that are severely corrupted by artefacts (blood vessels) or noise (Sec. IV).

Although we do not claim to achieve the state-of-the-art in these different examples, the fact that we use very simple processing (threshold of the Fourier-Argand response map) and readily achieve excellent results is an indication that the Fourier-Argand representation could be used as a building block of significantly more sophisticated algorithms, with good prospect to build competitive edge detection, pattern matching algorithms, in particular in the case of large noise.

II. FOURIER-ARGAND REPRESENTATION

Throughout the study presented in this paper, the following intuition plays an important role: when 2D rotations are involved, the complex representation $x + iy$ is more suitable than the Cartesian representation (x, y) , in the sense that the complex representation links the geometric operation — rotation by an angle α — to a simple algebraic operation — multiplication by a constant $e^{i\alpha}$. A function $h(x, y)$ can then be seen as a periodic function of the phase of the complex number $x + iy$. From this perspective, we construct, in the following sections, an optimal rotation-covariant representation. We name this representation the Fourier-Argand representation.¹

A. Approximation of Rotated Patterns

To have a better understanding of the steerable convolution problem, a natural question arises as follows: *For any given pattern (or filter) h and any positive integer N , what is the optimal basis consisting of N functions for approximating all rotated versions of h ?*

In order to address this problem, we first need to quantify the quality of an approximation. Let $\{\varphi_1, \varphi_2, \dots, \varphi_N\}$ be an arbitrary basis of N filters. Assume the approximation error is measured by the L^2 -distance between h and its approximation. Then, the best linear combination of those N basis functions that yields the minimum approximation error is the orthogonal projection $\mathcal{A}\{h\}$ of h onto the approximation space $\text{span}\{\varphi_1, \varphi_2, \dots, \varphi_N\}$. To find the optimal basis for approximating all rotated versions of h , we need to minimise the average approximation error e_N defined as follows,

$$e_N^2 \stackrel{\text{def}}{=} \frac{1}{2\pi} \int_0^{2\pi} \|a_h - \mathcal{A}\{a_h\}\|^2 d\alpha. \quad (5)$$

Without loss of generality, we can restrict the basis to be orthonormal, in which case the projection operator can be written as

$$\mathcal{A}\{f\} = \sum_{n=1}^N \langle f, \varphi_n \rangle \varphi_n, \quad (6)$$

¹Jean-Robert Argand is credited for the geometric interpretation of complex numbers $x + iy$ at the beginning of the 19-th century [18].

where $\langle f, g \rangle \stackrel{\text{def}}{=} \iint f(x, y)g^*(x, y) dx dy$, and $*$ denotes complex conjugation..

Notice that we make no other assumption on the basis, so that the formulation covers all possible linear representations. Surprisingly, solely minimising the average approximation error (5) automatically leads to a representation that has, additionally, the desired properties: rotation-covariant basis functions; and, simple relation between the linear coefficients and the angle of rotation (See Sec. II-C). We call this representation “Fourier-Argand”, because it uses the rotation-covariant complex variable $x + iy$ (Argand) and of the Fourier variable $e^{i\alpha}$, where i is the imaginary unit such that $i^2 = -1$.

B. Definition of the Fourier-Argand Representation

For a square integrable function $h(x, y)$, its Fourier-Argand representation is defined as the following power series in $\frac{x+iy}{|x+iy|}$,

$$h(x, y) = \sum_{k \in \mathbb{Z}} h_k(|x + iy|) \left(\frac{x + iy}{|x + iy|} \right)^k = \sum_{k \in \mathbb{Z}} H_k(x, y), \quad (7)$$

where \mathbb{Z} denotes the set of all integers. The coefficients h_k of the power series are functions of the modulus $r = |x + iy|$, and can be computed as follows,

$$\begin{aligned} h_k(r) &= \frac{1}{2\pi} \int_0^{2\pi} h(r \cos \theta, r \sin \theta) e^{-ik\theta} d\theta \\ &= \frac{r^k}{2\pi i} \oint_{C(r)} h(x, y) z^{-k-1} dz, \end{aligned} \quad (8)$$

where $z = x + iy$. Expressed in the polar coordinates, this equation is nothing but the Fourier series decomposition of $h(r \cos \theta, r \sin \theta)$ over θ . The contour $C(r)$ is the (counter-clockwise) circle of radius r centred at the origin. We refer to the functions $H_k(x, y)$ as Fourier-Argand basis functions. We also call the transform defined in (8), which maps a function $h(x, y)$ of two variables to a series of functions $h_k(r)$ of one variable, the Fourier-Argand transform.

The representation (7) is valid for complex-valued functions h . If h is real-valued, then the basis functions H_k has the Hermitian symmetry $H_{-k} = H_k^*$, $\forall k \in \mathbb{Z}$, where $*$ denotes complex conjugation.

C. Properties of the Fourier-Argand Representation

1) *Rotation Covariance*: One important property of the Fourier-Argand representation is that the basis functions $H_k(x, y)$ are rotation-covariant, defined as:

$${}^a H_k(x, y) = H_k(x, y) e^{-ika}, \quad (9)$$

for all $k \in \mathbb{Z}$ and $\alpha \in [0, 2\pi)$ [19]. The rotation-covariance shows that the Fourier-Argand functions H_k form a complete basis of the space consisting of all rotated versions of h . More importantly, these basis functions are eigenvectors of the rotation operator, which can be considered a stronger version of the steerability. On the one hand, it implies that the

approximation space is rotation-invariant, i.e., for every non-negative integer K ,

$$\left(\sum_{k=-K}^K H_k \right) (x, y) = \sum_{k=-K}^K H_k(x, y) e^{-ik\alpha}. \quad (10)$$

Note that the simple relation between the angle of rotation α and the linear coefficient not only makes it easy to rotate the filter h , but also makes the inverse problem — retrieving α from the coefficients — much easier to solve and robust to noise. We will demonstrate later in Sec. III how to estimate the local direction and how much accuracy and robustness to noise the estimate yields. Besides, the rotation-covariance is the essential reason for the optimality of the Fourier-Argand representation as we will show in the next section.

2) *Optimality*: As discussed in the Introduction, many representations have the steerability property. What makes the Fourier-Argand representation of particular interest is the fact that it is the optimal linear representation in terms of the approximation quality, which is stated by the following theorem.

Theorem 1: For any square integrable function $h(x, y)$, denote $\lambda_k \stackrel{\text{def}}{=} \langle H_k, H_k \rangle \geq 0$ for all integers k , where $H_k(x, y)$ are the Fourier-Argand functions defined in (7). Then, for any positive integer N , the optimal orthonormal basis $\{\varphi_1, \varphi_2, \dots, \varphi_N\}$, i.e., the orthonormal basis that minimises the average L^2 approximation error (5), is given by

$$\varphi_n(x, y) = \frac{1}{\sqrt{\lambda_{k_n}}} H_{k_n}(x, y), \quad n = 1, 2, \dots, N, \quad (11)$$

where the integers k_1, k_2, \dots, k_N can be positive or negative, and are chosen such that

$$\lambda_k \leq \lambda_{k_n}, \quad \forall n = 1, 2, \dots, N, \quad \text{where } k \notin \{k_1, k_2, \dots, k_N\}.$$

Moreover, the minimum approximation error achieved is given by

$$e_N^2 = \|h\|^2 - \sum_{n=1}^N \lambda_{k_n} = \|h\|^2 - \sum_{n=1}^N \langle H_{k_n}, H_{k_n} \rangle. \quad (12)$$

The proof of the theorem (a form of analytic principal component analysis) can be found in Appendix A, in which the rotation-covariance of the Fourier-Argand functions plays the central role.

The theorem shows that among all possible linear representations, the Fourier-Argand representation is the best for approximating all rotated versions of an arbitrary pattern h , which is in general not steerable. Notice that in addition to the rotation-covariance of the Fourier-Argand representation, its steerability ensures that the error of the approximation to any rotated version of the target pattern h is not dependant on the angle of rotation.

As a corollary, the rotation-covariant Fourier-Argand approximation requires the smallest number of basis functions for a fixed approximation error — indeed, significantly fewer basis functions are needed than using the windowed polynomial approximation (4) as suggested in [3]. This is far more than an improvement on computational efficiency, since, in practice, we cannot use polynomials of very high degree

due to their numerical instability. Therefore, as discussed in the Introduction, the Fourier-Argand representation provides a methodology for handling rotations, by extending the idea of steerable linear expansion from simple filters to general 2D patterns, which used to be difficult to handle.

D. Computation of the Fourier-Argand Representation

We show here that there are efficient ways to conduct the Fourier-Argand decomposition both analytically and numerically.

1) *Analytic Computation*: If we have the analytic formula of the target function h , we can construct its Fourier-Argand representation analytically, in some cases of interest.

Firstly, the most direct way is to compute the Fourier-Argand transform through its definition (8). For instance, we will show later in Sec. II-E how to compute the Fourier-Argand transform of elongated 2D Gaussian functions, i.e., Gaussian functions that spread differently in x - and y -direction, using formula (8).

Secondly, if h can be written as a product of two functions, i.e., $h(x, y) = g(x, y)f(x, y)$, and we know the Fourier-Argand representation of both g and f , then the Fourier-Argand representation of h are easily obtained by taking the product of the two series. Let the Fourier-Argand transform of g and f be g_k and f_k , respectively, as defined in (8). Then the Fourier-Argand transform of h is given by the discrete convolution between g_k and f_k , i.e.,

$$h_k(r) = \sum_{l \in \mathbb{Z}} g_l(r) f_{k-l}(r). \quad (13)$$

This is particularly useful when one of the two factors, say g , has a finite expansion, since the convolution in (13) becomes a finite sum. For example, the x -derivative (or any other partial derivative) of a Gaussian function (either elongated or not) is a product of a bivariate polynomial and the Gaussian function itself. The Fourier-Argand representation of a polynomial has finite terms. Therefore, the Fourier-Argand representation of functions of the form $P(x, y)G(x, y)$, where P denotes a bivariate polynomial and G is an arbitrary bivariate Gaussian function, can be computed easily. As a special case, the Fourier-Argand representation of the steerable filters discussed in [3], [4] can all be computed analytically.

Thirdly, one can compute the Fourier-Argand representation either in the space domain or in the frequency domain. This is a consequence of the fact that for any given target function h , the Fourier-Argand function $H_k(x, y)$ and its 2D Fourier transform mathematically share the same angular part, while their radial parts are a Hankel pair. To be clear, recall that the Fourier-Argand function of index k is

$$H_k(x, y) = h_k(|x + iy|) \left(\frac{x + iy}{|x + iy|} \right)^k.$$

One can verify that the 2D Fourier transform of $H_k(x, y)$ is given by

$$\hat{H}_k(\omega_x, \omega_y) = 2\pi (-i)^k \mathcal{H}_k\{h_k\} (|\omega_x + i\omega_y|) \left(\frac{\omega_x + i\omega_y}{|\omega_x + i\omega_y|} \right)^k, \quad (14)$$

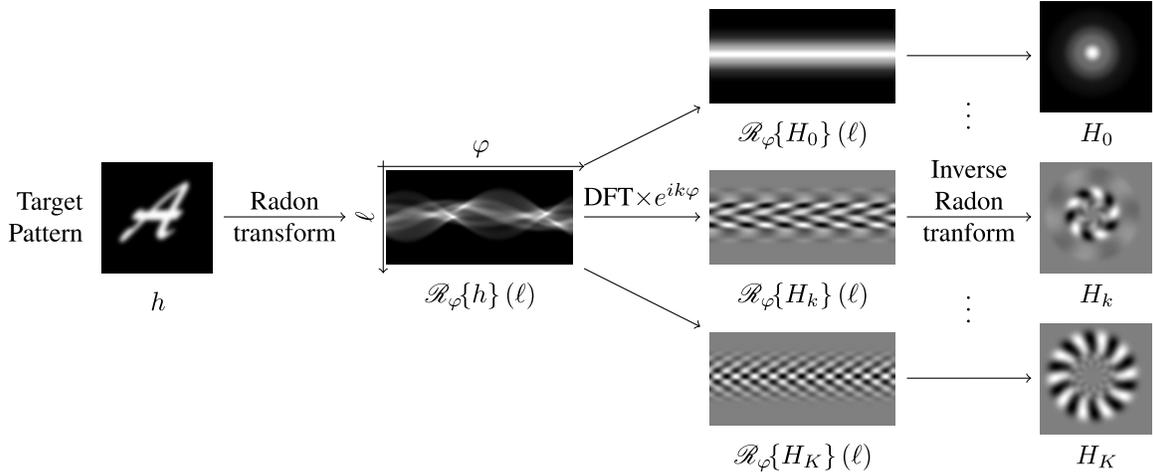


Fig. 1. Numerical computation of the Fourier-Argand representation using discrete Radon transform: 1. Discrete Radon transform of the pattern h yields a sinogram $\mathcal{R}_\varphi\{h\}(\ell)$. 2. DFT of the sinogram then multiply by $e^{ik\varphi}$. 3. Inverse Radon transform recovers the Fourier-Argand basis functions H_k .

where $\mathcal{H}_k\{\cdot\}$ denotes the Hankel transform of order k , which maps a univariate function to another univariate function, as defined below,

$$\mathcal{H}_k\{f\}(\rho) = \int_0^\infty f(r) J_k(\rho r) r dr, \quad (15)$$

where J_k is the k -th order Bessel function of the first kind. As an example, for any $a, b > 0$, it is not easy to compute the Fourier-Argand transform of $\exp(-a|x| - b|y|)$. However, finding the Fourier-Argand transform of its 2D Fourier transform $\frac{4ab}{(\omega_x^2 + a^2)(\omega_y^2 + b^2)}$ is relatively simple.

Finally, a closer look at the Fourier-Argand representation unveils a useful link with the Radon transform [20], which is extensively studied in the field of computed tomography. Classically, the Radon transform of a bivariate function h is defined by,

$$\mathcal{R}_\varphi\{h\}(\ell) = \iint h(x, y) \delta(x \cos \phi + y \sin \phi - \ell) dx dy, \quad (16)$$

where $\ell \in \mathbb{R}$ and $\varphi \in [0, 2\pi)$. The rotation-covariance of the Fourier-Argand basis H_k implies that its Radon transform has the following property:

$$\mathcal{R}_\varphi\{H_k\} = \mathcal{R}_0\{H_k\} e^{ik\varphi}. \quad (17)$$

As a consequence, the Radon transform of h becomes the following Fourier series in φ ,

$$\mathcal{R}_\varphi\{h\} = \sum_{k \in \mathbb{Z}} \mathcal{R}_0\{H_k\} e^{ik\varphi}, \quad (18)$$

which shows that $\mathcal{R}_0\{H_k\}$ is given by the Fourier integral

$$\mathcal{R}_0\{H_k\} = \frac{1}{2\pi} \int_0^{2\pi} \mathcal{R}_\varphi\{h\} e^{-ik\varphi} d\varphi. \quad (19)$$

(17-19) shows how to compute the Fourier-Argand basis H_k for every k :

1. Compute the Radon transform of h ($\rightsquigarrow \mathcal{R}_\varphi\{h\}(\ell)$).
2. Compute the Fourier series decomposition of $\mathcal{R}_\varphi\{h\}(\ell)$ with respect to φ ($\rightsquigarrow \mathcal{R}_0\{H_k\}$).

3. Multiply the coefficients found in step 2 by $\exp(ik\varphi)$ ($\rightsquigarrow \mathcal{R}_\varphi\{H_k\}$).
4. Perform the inverse Radon transform (filtered back-projection).

See Supplementary Material for some useful Fourier-Argand transforms.

2) *Numerical Computation*: In applications like pattern matching, the analytic expression for the pattern is usually unavailable, and a sampled version of the target function is all we have. In these cases, we need to compute the Fourier-Argand representation numerically.

The most straightforward numerical approach is to convert the digital pattern into polar coordinates through resampling. This method, however, is either slow or very inaccurate because sampling the polar coordinates uniformly leads to either very coarse resolution away from the centre (but reasonable speed), or adequate resolution everywhere but inefficient implementation (resolution at the centre is too fine).

Fortunately, the relation between the Fourier-Argand representation and the Radon transform discussed in the previous section leads to an efficient and relatively accurate algorithm for the numerical computation. See Fig. 1 for details.

If the analytic approach is applicable, each basis function is constructed explicitly from its expression. Otherwise, for a numerical pattern of M pixels, to obtain K basis functions, the computational cost is as follow: firstly, a discrete Radon transform of the input pattern costs $\mathcal{O}(LM)$, where L is the number of angles (around 90 in our experiments); next, a 1D Fourier transform applied to the rows of the Radon map costs $\mathcal{O}(\sqrt{M} L \log L)$; and finally, inverse Radon transforms applied to the K components $\mathcal{R}_\varphi\{H_k\}$ cost $\mathcal{O}(KLM)$. Note that the size of the pattern is usually much smaller than that of the target image. As a consequence, the computational cost to obtain the Fourier-Argand basis is significantly lower than that of other steps of the proposed algorithm.

E. Example

We illustrate the Fourier-Argand transform by considering an *elongated* Gaussian function. On the one hand, isotropic

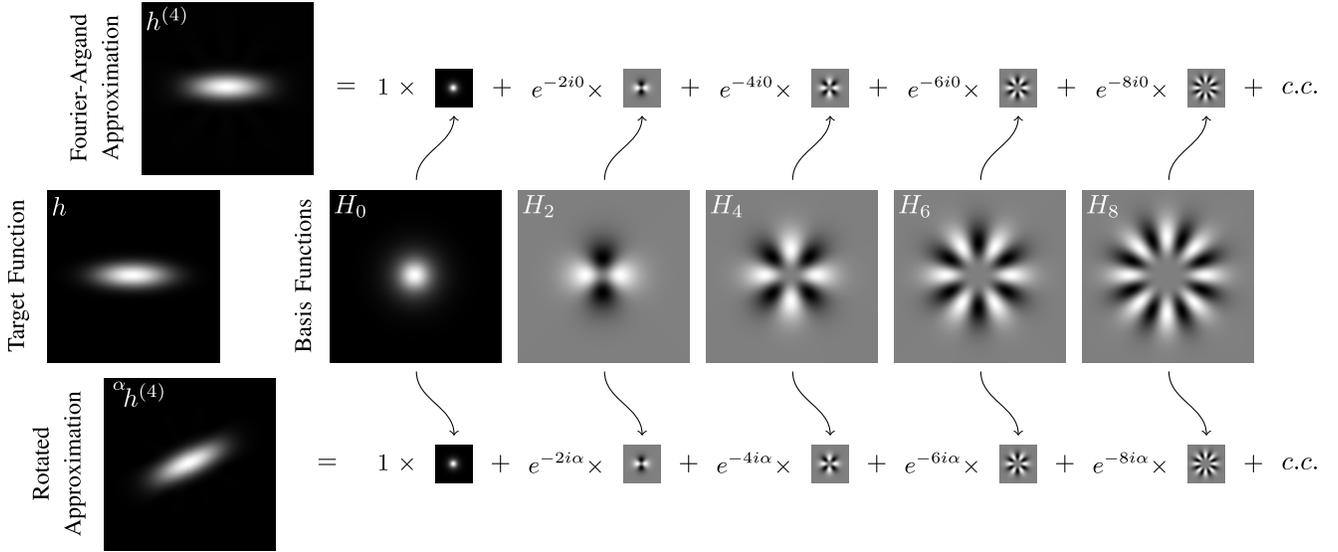


Fig. 2. Demonstration of the Fourier-Argand representation for an elongated Gaussian function h with ellipticity $\epsilon = \frac{2}{3}$. For better visualisation, each basis function H_{2k} is rescaled according to its norm $\|H_{2k}\|$, and only the real part is shown in the figure. These 5 basis functions, together with their complex conjugates (denoted by *c.c.* in the figure), sum up to the Fourier-Argand approximation $h^{(4)}$ of h . The approximation yields an L^2 -error of 2.19×10^{-2} (i.e., 33.2 dB SNR) with the norm of the basis function H_{2k} being 8.73×10^{-1} , 3.10×10^{-1} , 1.34×10^{-1} , 6.14×10^{-2} and 2.87×10^{-2} , for $k = 0, 1, 2, 3$ and 4. This fast decrease ensures that with very few basis, the Fourier-Argand approximation can already reach a high accuracy. The summation of the same basis functions multiplied by complex exponentials $e^{-i2k\alpha}$ form an approximation to the rotated Gaussian function ${}^a h$, respectively.

Gaussian functions, together with their derivatives, are successful filter models for both the detection and the extraction of local structures in images (e.g., ridges and edges). On the other hand, lacking necessary elongation, they have limited directional selectivity, which degrades their performance in various applications. In this section, We derive the Fourier-Argand representation of the elongated Gaussian function, and show that it is possible to approximate it with few terms, even in large ellipticity cases. We also compare the Fourier-Argand approximation with the widely used windowed polynomial approximation suggested by [3], [4] in terms of their efficiencies.

1) *Computation of the Fourier-Argand Representation:* Let h denote the elongated Gaussian function shaped by two parameters a and b , and defined as follows,

$$h(x, y) = \frac{1}{\pi ab} \exp\left(-\frac{x^2}{a^2} - \frac{y^2}{b^2}\right). \quad (20)$$

Noticing that the symmetry $h(x, y) = h(-x, -y)$ implies that the Fourier-Argand transform $h_k(r)$ vanishes for odd integer k , we rewrite the representation (7) as,

$$h(x, y) = \sum_{k \in \mathbb{Z}} \underbrace{h_{2k}(|x + iy|)}_{H_{2k}(x, y)} \left(\frac{x + iy}{|x + iy|}\right)^{2k}. \quad (21)$$

The non-vanishing radial functions h_{2k} are directly obtained from the Fourier integral (8), i.e.,

$$\begin{aligned} h_{2k}(r) &= \frac{1}{2\pi^2 ab} \int_{-\pi}^{\pi} e^{-\frac{r^2 \cos^2 \theta}{a^2} - \frac{r^2 \sin^2 \theta}{b^2}} e^{-i2k\theta} d\theta \\ &= \frac{(-1)^k}{\pi ab} I_k\left(\frac{1}{2}r^2(a^{-2} - b^{-2})\right) e^{-\frac{1}{2}r^2(a^{-2} + b^{-2})}, \end{aligned} \quad (22)$$

where $I_k(\cdot)$ denotes the order- k modified Bessel function of the first kind.

The following partial sum of the series (21) is an optimal approximation of the elongated Gaussian function h :

$$h^{(K)}(x, y) := \sum_{k=-K}^K H_{2k}(x, y). \quad (23)$$

We call this partial sum the order- K Fourier-Argand approximation. The number of terms can be determined by setting an upper-bound on the approximation error. In a pattern matching problem, this error quantifies how well the pattern should be approximated; i.e., how discriminant its approximation is, compared to the other patterns in the image. For an elongated Gaussian, it can be shown that this number is directly related to the ellipticity

$$\epsilon = \frac{b - a}{a}.$$

See Supplementary Material for a detailed analysis.

We visualise, in Fig. 2, the order-4 Fourier-Argand approximation $h^{(4)}$ of an elongated Gaussian function with ellipticity $\epsilon = \frac{2}{3}$. The real part of the (normalised) basis functions are shown for $K = 4$. The basis functions H_{2k} sum up to the optimal approximation $h^{(4)}$. The norm of the basis function H_{2k} decreases quickly from 8.73×10^{-1} to 6.14×10^{-2} as k increases from 0 to 4, resulting in an L^2 approximation error of 2.19×10^{-2} (i.e., 33.2 dB SNR).

According to (10), $h^{(4)}$ can be rotated exactly by an arbitrary angle α :

$${}^a h^{(K)}(x, y) = \sum_{k=-K}^K H_{2k}(x, y) e^{-i2k\alpha}.$$

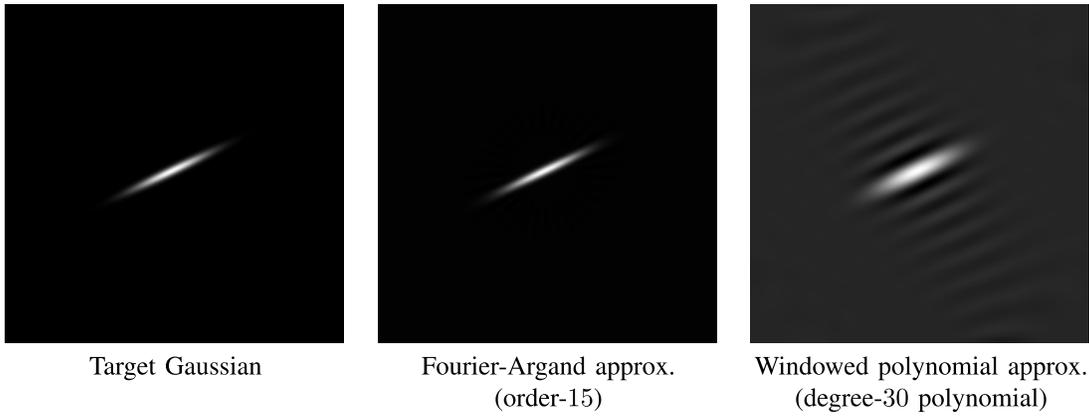


Fig. 3. Comparison of approximation quality. The target function being approximated (left) is an elongated Gaussian function with ellipticity 0.9. The order-15 Fourier-Argand approximation (middle) yields an approximation error of 2.40×10^{-2} (i.e., 32.4 dB SNR). The windowed polynomial approximation (right), which uses monomials of degree as high as 30, yields a much larger approximation error, 5.07×10^{-1} (i.e., 5.90 dB SNR). Both approximations use 31 basis functions, which demonstrates that the Fourier-Argand approximation is significantly more efficient than the windowed polynomial approximation. Note that the polynomials used here is already of a high degree, and in practice, it is difficult to handle polynomials of a higher degree.

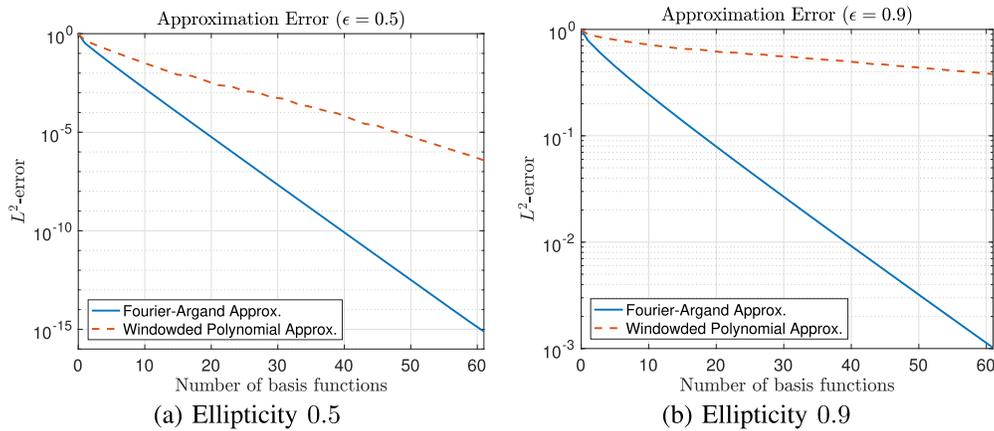


Fig. 4. The Fourier-Argand representation is significantly more efficient than usual steerable bases. The approximation errors, for both the Fourier-Argand approximation (solid line) and the windowed polynomial approximation (dashed line), are plotted against the number of basis functions involved. The target function under approximation can be either an elongated Gaussian (a) with ellipticity 0.5 or an extremely elongated Gaussian (b) with ellipticity 0.9. In both cases, the error decreases significantly faster for the Fourier-Argand approximation than that for the windowed polynomial approximation, as the number of basis functions increases.

We demonstrate this process in the bottom row of Fig. 2. Note that writing ${}^a h^{(K)}$ causes no ambiguity, since by construction we have $({}^a h)^{(K)} = {}^a(h^{(K)})$.

2) *Optimality of the Fourier-Argand Approximation:* Compared to other steerable representations, the Fourier-Argand representation exhibits significantly higher efficiency. To demonstrate our claim, we compare the Fourier-Argand approximation with the successful and widely used windowed polynomial approximation, first proposed by Freeman *et al.* in [3]. In their paper, the authors proposed to use a bivariate polynomial multiplied by an isotropic window function for approximating the target function h . In the Gaussian case here, h is a real-valued function, and has the symmetry $h(x, y) = h(-x, -y)$. Therefore, the windowed polynomial approximation can be written as follows,

$$h_{PK}(x, y) = \sum_{k=0}^K \sum_{k'=0}^k a_{k,k'} x^{k'} y^{2k-k'} W\left(\sqrt{x^2 + y^2}\right), \quad (24)$$

where the window function W is the isotropic Gaussian,

$$W(r) = e^{-\frac{r^2}{\max(a,b)^2}},$$

$\max(a, b)$ is, empirically, the choice of standard deviation that yields the best approximation. Notice that due to the symmetry of h , (24) only contains the even terms of the polynomial, and the degree of the polynomial is $2K$ rather than K . With the above notations, both the windowed polynomial approximation h_{PK} , (24), and the Fourier-Argand approximation $h^{(K)}$, require the same number ($N = 2K + 1$) of real-valued basis functions.

As shown in Fig. 4, with the same number of basis functions, the Fourier-Argand approximation achieves a significantly lower error than the windowed polynomial approximation. As the number of basis functions increases, the approximation error decreases significantly faster for the Fourier-Argand approximation than that for the windowed polynomial approximation, especially when the target function under approximation is extremely elongated.

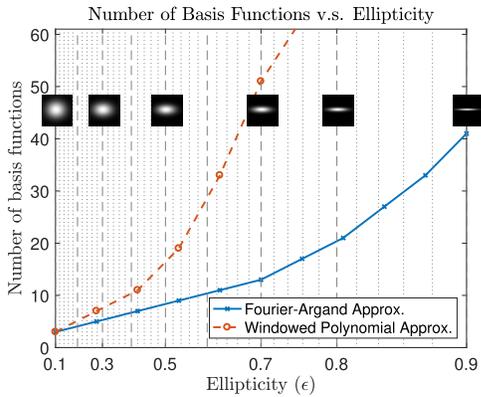


Fig. 5. The minimum number of basis functions needed by an approximation of L^2 -error not exceeding 10^{-2} (SNR = 40 dB) v.s. the ellipticity of the target Gaussian function.

Correlatively, comparing with the windowed polynomial approximation, the Fourier-Argand approximation requires significantly fewer basis functions to reach the same approximation error. In order to yield an L^2 -error below 10^{-2} (i.e., 40 dB SNR), the minimum number of basis functions needed increases as the target Gaussian function becomes more elongated. The windowed polynomial approximation experiences a much faster increase compared to the Fourier-Argand approximation. As a consequence, the windowed polynomial approximation quickly needs to deal with very high-degree polynomials, which is cost inefficient and numerically unstable. Fig. 3 demonstrates the approximation quality when the Gaussian function is extremely elongated ($\epsilon = 0.9$). We see that even if a degree-30 polynomial is used, the windowed polynomial approximation $h_{P_{15}}$ is still not satisfactory. Meanwhile, with the same number of basis functions, the Fourier-Argand approximation $h^{(15)}$ is much more accurate.

III. MAXIMISING LOCAL RESPONSE

As discussed in the Introduction, once we have a rotation-covariant representation, we can effectively retrieve the local orientation of the filter/pattern in the image. Formulated in a general matched filtering framework, the problem involves convolving the image $I(x, y)$ with a filter $h(x, y)$, while allowing the filter to be rotated by an arbitrary angle α . The local direction is then given by the angle that maximises the filter response at each location. Mathematically, this steerable matched filtering problem is formulated as follows,

$$R(x, y) = \max_{\alpha} (I * {}^{\alpha}h)(x, y), \quad (25)$$

where $R(x, y)$ is the final filter response, and the angle $\alpha^*(x, y)$ that achieves the maximum is called the local orientation (or direction) of the image I at (x, y) . This is a quite general framework. For instance, we can take a simple 2D function as a filter to detect image local structures, such as ridges, edges and wedges. The Canny-like filters and the steerable filters all share this formulation [4]. We can also consider a particular pattern as a filter, in which case (25) can

also be used for pattern recognition. In both cases, being able to accurately recover the local direction is crucial.

As discussed in the Introduction, the naive approach to solve (25) relies on a large number of filtering operations. Due to the rotation-covariance of the Fourier-Argand representation (7) and (9), and the linearity of the convolution operator, the problem (25) is reduced to,

$$\alpha^*(x, y) = \arg \max_{\alpha} \sum_{k \in \mathbb{Z}} (I * H_k)(x, y) e^{-ika}, \quad (26)$$

which is much easier to solve when the above sum only contains a few number of terms (Fourier-Argand approximation). In this section, we show in details how to estimate the local orientation with the Fourier-Argand representation of the filter. Throughout this section, we assume that the image I and the target function h are real-valued.

A. The Frequency Estimation Algorithm

According to the discussion in Sec. II, we construct the Fourier-Argand approximation of the target filter h , which yields a controllable small approximation error, i.e.,

$$h^{(K)}(x, y) = \sum_{k=-K}^K H_k(x, y). \quad (27)$$

Next, we convolve of the image $I(x, y)$ with the basis functions $H_k(x, y)$ for $0 \leq k \leq K$, i.e.,

$$I_k(x, y) = I(x, y) * H_k(x, y). \quad (28)$$

Then (26) becomes

$$\alpha^*(x, y) = \arg \max_{\alpha} \sum_{k=-K}^K I_k(x, y) e^{-ika}. \quad (29)$$

This problem can be solved exactly by setting to zero the derivative of the objective in (29) with respect to α . The maximisation problem is then solved by finding the roots on the unit circle of the following polynomial

$$P(z) = \sum_{k=-K}^K -ik I_k(x, y) z^k. \quad (30)$$

Notice that the symmetry $P(z) = P^*(z^{-1})$ ensures that if z_0 is a root of P then z_0^{-1} is also a root of P . Therefore, finding the roots of P eventually amounts to finding the roots of a polynomial of degree K (instead of $2K$).

Alternatively, we can also solve (29) approximately, by maximising the real part of an L -point Discrete Fourier Transform (DFT) of the sequence $(\frac{1}{2}I_0, I_1, I_2, \dots, I_K)$. This numerical algorithm provides an angle accuracy of $\frac{\pi}{L}$. For instance, if we choose $L = 360$, then the angle accuracy is 0.5° , which is usually sufficient in practice. If in some extreme cases where higher angle accuracy is required, the approximate solution provided by the DFT method can always serve as a good initialisation for a numerical method, e.g., Newton's method, finding the roots of (30) within an arbitrary error.

The above approximate algorithm is very fast. We tested the MATLAB implementation on a laptop with a dual

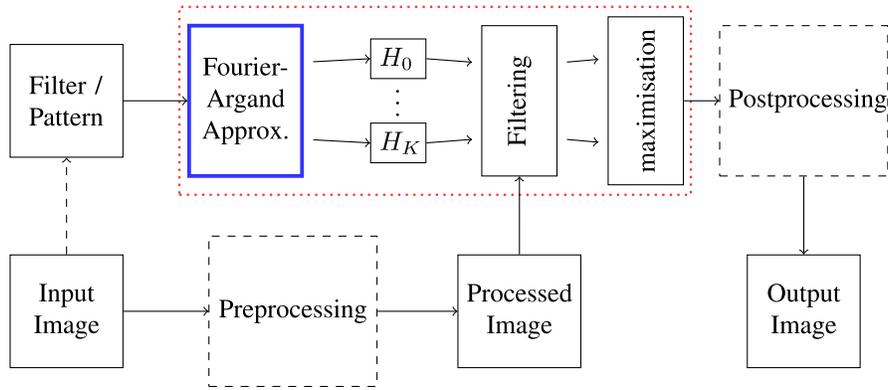


Fig. 6. Overall pipeline. The Fourier-Argand approximation (thick blue box) can be constructed either analytically or numerically as detailed in Fig. 1. Dashed boxes denote optional processings. The dashed arrow means the pattern can simply be an image patch containing the pattern of interest. Key processings of the proposed algorithm are enclosed in the dotted red box.

core 2.6 GHz Intel i5 CPU and 8 GB memory. For an 512×512 grey-scale image, and a Fourier-Argand approximation of 41 basis functions, the filtering process takes less than 2 seconds, while maximising the response requires less than 3 seconds, resulting in a total time consumption of less than 5 seconds. As a comparison, to reach the same angle accuracy (i.e., 0.5°) without the Fourier-Argand representation, one needs to filter the image with 360 rotated versions of the filter, which is several times more time consuming.

The overall pipeline of the proposed framework for steerable pattern/filter matching is summarised in Fig. 6.

B. Demonstration

The algorithms discussed in previous sections provide an efficient method for steerable matching. We now demonstrate the algorithms by an example, where we try to find different rotations of a pattern in an image. The goal of this demonstration is not to design a delicate pattern recognition algorithm. Instead, we show here that with the help of the proposed Fourier-Argand representation, simple processing (e.g., filtering, FFT) suffices to produce a reasonably good and fast (within seconds) detection result. Indeed, the algorithm also gives extra information — accurate orientation of the patterns.

Rotation invariance has played an important role for the success of many pattern recognition algorithms. To achieve rotation-invariance, one common practice is to discretise the angle of rotation and build various statistics, e.g., histograms, according to the discrete orientations [21]. Instead of discretising the angle of rotation, other invariants exploit moments that are rotation-covariant [13], [14], [16]. The magnitude of these moments is invariant to rotation, while the phase varies as the orientation of the pattern changes. However, due to the lack of a systematic way to incorporate the phase information into the descriptors [14], the phase is usually discarded. However, exploiting the phase information is of crucial importance for an accurate recognition, since the local orientation of the pattern is purely encoded in the phase of the moments.

Consider the problem of detecting Hong Kong 2-dollar coins in a real-scene image captured by a standard mobile phone

camera (Fig. 7). We are only interested in finding the 2-dollar coins with their obverse side up, and want to distinguish them from both the same 2-dollar coins, but with their reverse side up, and coins of other types.

We first take the magnitude of the Laplacian of the input image — a simple high-pass filtering, to mitigate the obvious illumination changes between the coins. As for the template, we simply take the image patch that contains one of the 2-dollar coins. The pattern is taken from the high-pass filtered image, and marked by the red circle in Fig. 7(top right). As discussed in Sec. II-D.2 and illustrated by Fig. 1, we compute the Fourier-Argand approximation ($K = 20$, i.e., 41 basis filters) of the pattern using the discrete Radon transform. With this accurate approximation, the following matched filtering problem is efficiently (within seconds) solved by the algorithm introduced in Sec. III-A, where we get both the response $R(x, y)$ and the local orientation $\alpha^*(x, y)$:

$$R(x, y) = \max_{\alpha} \left(|\nabla^2 I| * {}^{\alpha} h^{(K)} \right) (x, y), \quad (31)$$

where $h^{(K)}$ is the Fourier-Argand approximation of the pattern, and $|\nabla^2 I|$ denotes the magnitude of the Laplacian of the input image I . The maximum is reached by $\alpha^*(x, y)$ at each location (x, y) . To further mitigate the illumination changes, it is standard to use the normalised cross-correlation between the pattern and the signal [23]. We compute the normalised cross-correlation between the Fourier-Argand approximation and the high-pass filtered image patch inside each disk $\mathcal{D}_{(x,y)}$, which is centred at (x, y) and share the same radius as the pattern, i.e.,

$$C(x, y) = \frac{R(x, y) - \mu_i(x, y)\mu_h}{\sigma_i(x, y)\sigma_h}, \quad (32)$$

where μ_i and σ_i denote the mean and standard deviation of the high-pass filtered image $|\nabla^2 I|$ inside the disk $\mathcal{D}_{(x,y)}$, respectively, and μ_h and σ_h denote the mean and standard deviation of $h^{(K)}$, respectively.

As shown in Fig. 7, the normalised cross-correlation (bottom left) has very sharp and distinguishable peaks, which collocate with the 2-dollar coin patterns. As a consequence, the occurrences of the pattern are reliably detected at the local

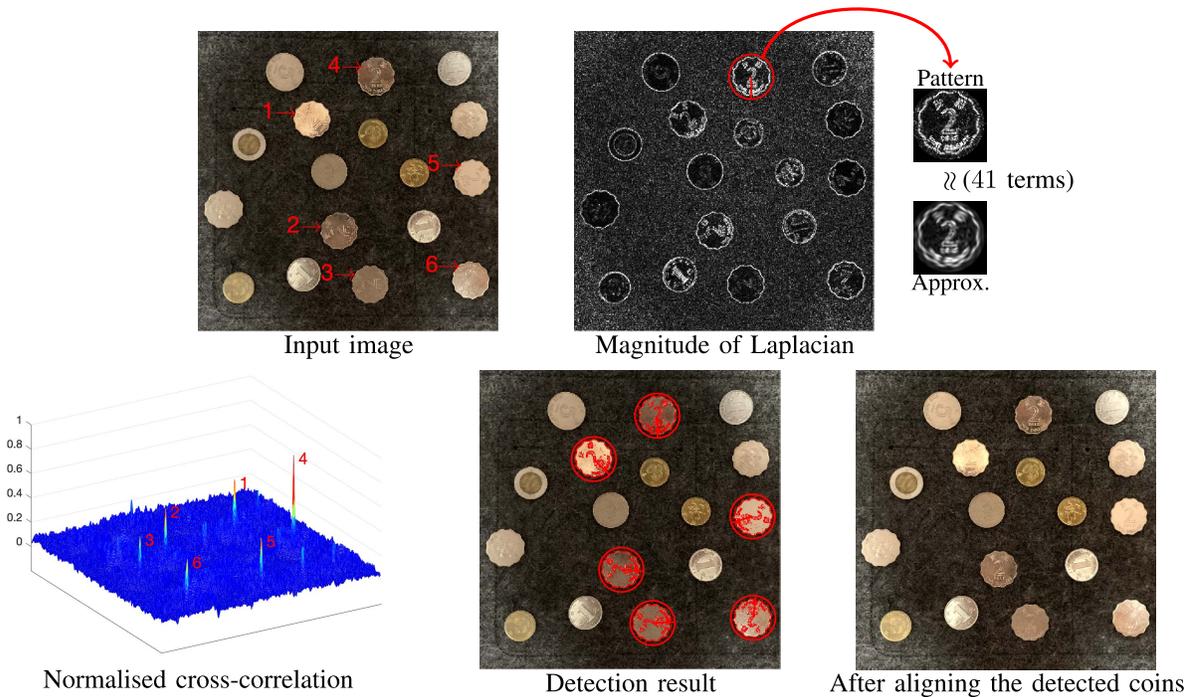


Fig. 7. Demonstration of real-scene pattern recognition. The 2-dollar coins are sought in a real-scene image. The image undergoes a high-pass filtering before the Fourier-Argand algorithm is applied. Maximised over all possible rotations, the normalised cross-correlation between the pattern and the high-pass filtered image exhibits sharp and large peaks. A simple threshold on the correlation map identifies the 2-dollar coins with very high location accuracy. The angle at which the maximum is reached gives an accurate estimate of the orientation for each coin, as shown by rotating the detected coins back using the estimated angles (bottom right picture). The overall computation time is less than 7 seconds.

maxima of the normalised cross-correlation (32). In the figure, each detected pattern is marked by a red circle. For a better visualisation, we create a binary mask by applying a threshold to the intensity of the target pattern, and overlay the mask at locations where the pattern is detected (the mask is also rotated to the estimated orientation). In spite of their changing appearance, all of the 2-dollar coins are successfully detected with very high location and angle accuracy. Also note that the algorithm correctly avoids those 2-dollar coins with their reverse side up. To further demonstrate the angle accuracy, we rotate each 2-dollar coin in the original image to its vertical position according to the estimated orientation. We can see that all of the obverse-sided 2-dollar coins are aligned very well with each other.

IV. EXPERIMENTS

The demonstration in the previous section shows that in normal conditions (moderate noise and normal lighting), the proposed method produces sharp responses — accurate in location and orientation. To further validate the proposed method, we conducted experiments on real images and synthetic images with very large noise.

We stress here that the goal of these experiments is not to show that the proposed algorithm is able to outperform the state-of-the-art algorithms on each of the tasks. Instead, the experiments show that with very simple added processing, the Fourier-Argand representation can be used to reach a reasonably good performance that is comparable with (if not better than) algorithms specifically designed for the tasks.

Therefore, it is reasonable to expect that with more careful design for a specific task, the proposed methodology would help increase the performance.

A. Retinal Blood Vessel Segmentation

Segmentation of the vasculature in retinal fundus images is very important for diagnosis and monitor of progressive disease [22]. In the green channel of the retinal images, blood vessels typically appear as curly dark lines. In spite of the low contrast and relatively high noise of the image, the difficulty in detecting and segmenting blood vessels owes much to the fact that the vessels spread in arbitrary directions.

Researchers in the medical imaging society usually solve this problem by using a filter bank that contains lines of several (typically 12) uniformly distributed orientations, and often adopt complicated pre- and post-processing [24], or rely on supervised learning [22].

We show by this experiment that using simple ridge filters and the proposed Fourier-Argand method, a simple thresholding on the response map readily gives comparable results with the algorithms relying on significantly more sophisticated processing and training [22].

We used a very simple scheme: 1) take the inverted green channel of the retinal image, for its good contrast between blood vessels and other tissues; 2) apply the Fourier-Argand filter (elongated Gaussian, $a = 1$ and $b = 10$); 3) subtract the local mean, i.e., the green channel filtered by a Gaussian with $\sigma = 10/\sqrt{2}$; 4) threshold the response (the threshold was set to be inversely proportional to the green channel).

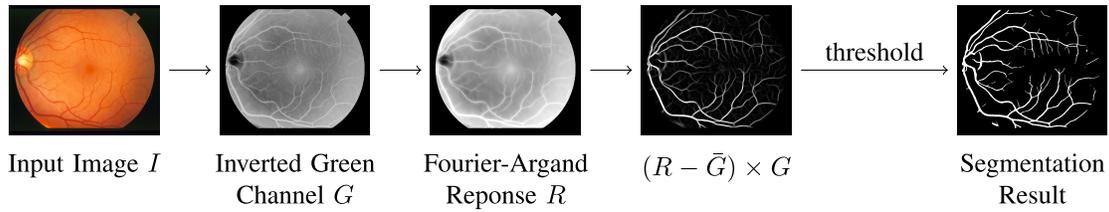


Fig. 8. Flowchart of blood vessel detection in retinal images using Fourier-Argand filter (elongated Gaussian $a = 1$, $b = 10$). On the STARE dataset, the AUC for such a simple scheme (quantifying the compromise between false and true detected pixels) reaches 0.9681, similar to the AUC that results from the much more sophisticated processing [22] (Line Detectors + SVM). Replacing Fourier-Argand filter by steerable filters results in a poorer AUC (0.9509). \bar{G} denotes the local mean of the green channel obtained by a Gaussian filtering.

TABLE II
PERFORMANCE ON STARE DATASET

Algorithm	AUC	Accuracy
Fourier-Argand Filter (this work)	0.9681	0.9614
Steerable Filter [4]	0.9509	0.9522
Fourier-Argand Moment [17]	0.9449	0.9483
Line Detectors + SVM [22]	0.9680	0.9646

The boundary of the field of view (FOV) was extended by filling the exterior with the local mean near the boundary. This is a standard processing to eliminate boundary effect. See Fig. 8 for a flowchart of this scheme.

To compare with other algorithms, the same scheme was used except that we replaced the response of Fourier-Argand filter by the steerable filter [4] (order 4, $\sigma = \sqrt{5}$) and the Fourier-Argand moment ridge detector ($\sigma = 10/\sqrt{2}$, $\sigma_c = 1$) [17]. The parameters were chosen to be empirically the best.

We tested the algorithms on the STARE dataset (20 images) [24]. The area under the receiver operating characteristic curve (AUC) was calculated (See Table II), which is a key measurement of the quality of the response map. We see that using a very simple processing, the proposed Fourier-Argand filter readily produced comparable results with the more sophisticated algorithm [22], which is a combination of several line detectors and supervised training (SVM). The AUC dropped to 0.9509 and 0.9449, when we replaced the Fourier-Argand filter by steerable filter [4] and Fourier-Argand moment [17], respectively.

B. Ridge Detection From Super Noisy Images

Ridges are commonly encountered when a high-pass filter is applied to a signal with singularities. For instance, in a high-pass band of a wavelet decomposition, ridges correspond to the edges of the original image. The ability to distinguish ridge pixels from non-ridge pixels gives the possibility to apply different operations on pixels in smooth regions of the image and in regions with discontinuities. This is very important for a better representation of the image, and is crucial for problems such as image de-noising [25], de-convolution [26], [27], etc. However, detecting ridges in the high-pass band is difficult due to the very poor SNR. The key to suppressing the noise for a successful ridge detection is to exploit the consistency of

the signal along the ridge, which requires accurate estimation of the local ridge directions and the use of highly direction-selective filters.

We have shown in Sec. II that a small number of Fourier-Argand basis functions are sufficient to form a very accurate approximation of extremely elongated Gaussian functions. Large ellipticity makes the Gaussian filter highly direction-selective, which is a property difficult to achieve by the steerable filter with windowed polynomial approximation [4].

To demonstrate the efficiency of the Fourier-Argand filters, we conducted an experiment on synthetic data. We generated 10 random curves with self-intersections and varying curvature. The images were corrupted by additive white Gaussian noise of different levels, corresponding to a PSNR ranging from -5 dB to 5 dB. At each noise level, 30 realisations of the random noise were generated. The overall measures were averaged over all noise realisations and all images.

Elongated Gaussian filters of ellipticity 0.9 ($b = 10a$) were used to detect the ridge pixels. It is easy to see that a^2 is proportional to the effective size of the filter. Therefore, we set a to be proportional to the standard deviation of the noise. At -5 dB, a is set to 3. Besides, we used the degree-10 Fourier-Argand approximation in the experiment.

We compare the proposed algorithm with the steerable filter algorithm [4], as well as the Fourier-Argand moment ridge detector [17]. We choose the order-4 steerable ridge detector, which uses the most direction-selective filter provided by the authors. We set the standard deviation of the isotropic Gaussian window used by the steerable filter to $\sigma = \sqrt{ab}/2$ pixels. For the Fourier-Argand filter, we set the window size σ to $b/\sqrt{2}$ and the ridge width σ_c to $a/\sqrt{2}$. These settings make sure that the algorithms work at the same scale, i.e. with the same spatial resolution, and is empirically the best choice.

Fig. 9 shows part of the ridge detection results, where the input noisy images are of very poor quality (PSNR -5 dB, and we can hardly see any detail with our bare eyes). All of the response maps of the three algorithms revealed the underlining signal to some extent. But we can see that the Fourier-Argand filter produced much more faithful — high contrast and better connectivity — response maps. The detection results were obtained by applying a non-maximum suppression to the response maps along the line perpendicular to the estimated local ridge direction, and followed by a hard thresholding. The threshold was chosen so that 0.5% of the pixels were kept. In Fig. 9, we marked the true-positives by

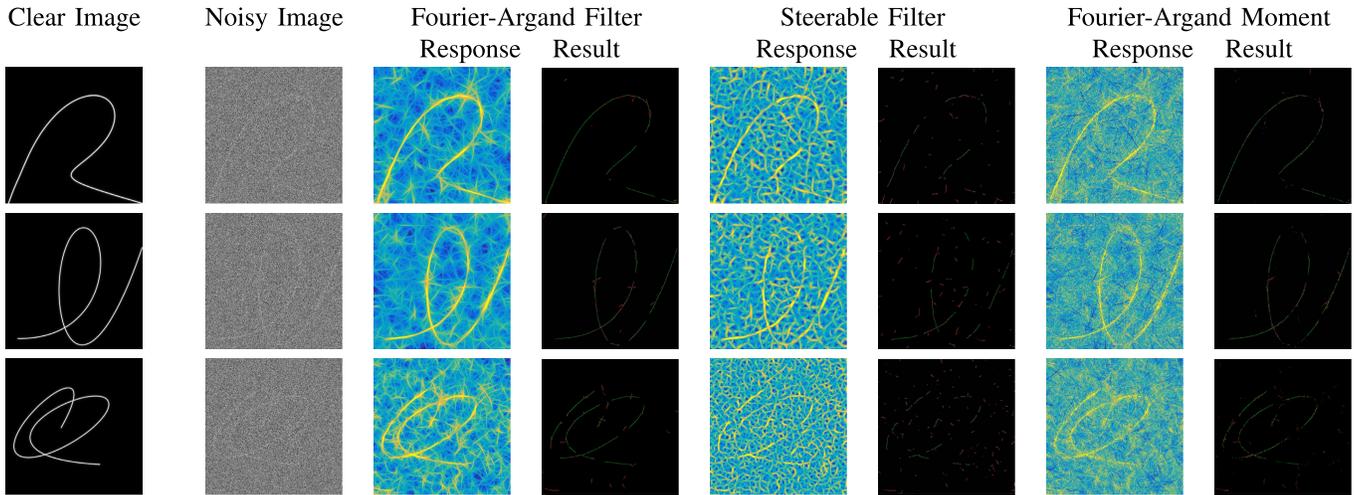


Fig. 9. Ridge detection from super noisy images. The clear input images were first corrupted by large white Gaussian noise, resulting in a PSNR of -5 dB. The response of the Fourier-Argand filter exhibits more contrast and is more continuous, and the corresponding detection result contains less false-positively (in red) and more true-positively (in green).

green points and the false-positives by red ones. We observed that the proposed Fourier-Argand filter usually produces more accurate estimation of the local direction, which benefits the non-maximum suppression. Overall, the Fourier-Argand filter achieved significantly better performance.

We also conducted quantitative assessment of the three algorithms (See Table III). Firstly, we considered a detected ridge pixel as “correct” (i.e. true-positive) if its Euclidean distance to the ridge curve is smaller than or equal to 1 pixel. Otherwise, the detected pixel was considered as false-positive. Notice that we are using a quite strict criterion, since some detected pixels close to the ridge (but not exactly within the 1-pixel range) are considered as false-positives. We then computed the precision (the number of true-positives over the number of detected pixels) and the recall (the number of true-positives over the number of ground truth ridge pixels). Besides, we computed the mean and median value of the distance between the detected pixels to the ridge curve, and the mean and median values of the angle error over all ground-truth ridge pixels.

Table III summarises the above measures for all of the three algorithms at 5 different noise levels. We can see that the Fourier-Argand filter constantly outperformed the other two algorithms, especially when the noise is very large. The steerable filter exhibited large location errors. This is consistent with the visual inspection in Fig. 9, the large number of false-positives dramatically decreased the location accuracy.

C. Pattern Matching Under Very Large Noise

Assisted by the high quality Fourier-Argand approximation and the frequency estimation algorithm, a steerable matched filtering is able to robustly detect patterns in the presence of large noise.

To demonstrate our claim, consider the problem of finding the letter ‘a’ in a synthetic 756×756 image containing multiple rotated versions of the five vowels, i.e., ‘a’, ‘e’, ‘i’, ‘o’ and ‘u’ (see Fig. 10, top left, all of the letters ‘a’ are shown in red).

TABLE III
QUANTITATIVE ASSESSMENT OF THE ALGORITHMS

Algorithm	Performance		Location Error		Angle Error	
	Rec.	Prec.	Med.	Mean	Med.	Mean
PSNR: 5 dB						
F.-A. Filter	0.88	0.97	0.28	0.37	0.92	1.39
Steerable Filter	0.82	0.91	0.33	6.06	3.27	4.02
F.-A. Moment	0.80	0.87	0.44	0.53	4.21	5.32
PSNR: 2.5 dB						
F.-A. Filter	0.88	0.96	0.31	0.41	0.87	1.42
Steerable Filter	0.76	0.83	0.41	10.30	3.50	4.29
F.-A. Moment	0.78	0.85	0.46	0.57	3.04	3.99
PSNR: 0 dB						
F.-A. Filter	0.82	0.91	0.35	0.65	0.90	1.58
Steerable Filter	0.62	0.67	0.61	17.21	3.90	4.81
F.-A. Moment	0.78	0.85	0.42	0.65	2.31	3.13
PSNR: -2.5 dB						
F.-A. Filter	0.74	0.81	0.47	1.50	1.00	1.85
Steerable Filter	0.42	0.46	1.32	28.27	4.37	5.56
F.-A. Moment	0.68	0.76	0.51	1.60	2.08	2.97
PSNR: -5 dB						
F.-A. Filter	0.56	0.62	0.75	5.31	1.18	2.34
Steerable Filter	0.24	0.26	9.65	39.52	5.30	6.75
F.-A. Moment	0.54	0.60	0.75	7.12	1.98	3.05

*. The unit for location error is “pixel”, and the unit for angle error is “degree”.

The image was then corrupted by a large additive Gaussian noise (PSNR = 0 dB, Fig. 10, top right). The pattern to be sought is an image patch containing a clear letter ‘a’. Through the same procedures that we used in Sec. III-B, the proposed algorithm robustly located all occurrences of the letter ‘a’ and

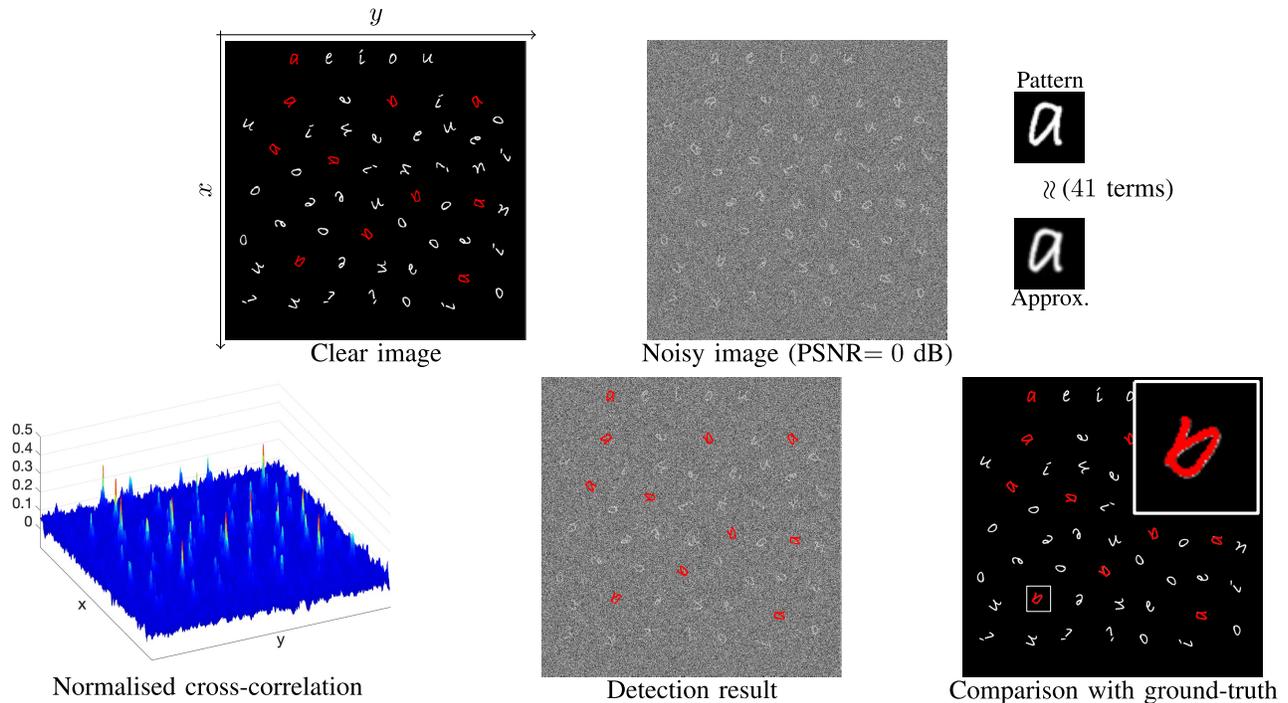


Fig. 10. Demonstration of pattern recognition under large noise. The letter ‘a’ is sought in an image containing multiple rotated versions of the five vowels (all letters ‘a’ are shown in red). The image is first corrupted by a large white Gaussian noise (PSNR = 0 dB). Searching in the noisy image with the high quality Fourier-Argand approximation, the proposed algorithm accurately identifies all rotated versions of the letter ‘a’. The large noise only causes slight error in the location and angle of the detect pattern. The algorithms requires less than 10 seconds to process this 756×756 image.

accurately retrieves their orientations. To visualise the result, the detected letters (shown in red) were overlaid on top of the noisy image in their estimated orientations (Fig. 10, bottom centre). To demonstrate both the location and angle accuracy of the algorithm, we also overlaid the detected letter ‘a’ on top of the noiseless image (Fig. 10, bottom right): a zoom over one of the detected patterns shows how accurate the estimated angle is, since most of the (ground-truth) noiseless pattern (white) is covered by the detected one (red).

The algorithm is very robust to noise. As we kept increasing the intensity of the noise, the algorithm still correctly detected about 82% (averaged over 100 noise realisations) of the letters when the PSNR drops to -5 dB, in which case we can hardly see any letter with our bare eyes.

Implemented using MATLAB on the commercial laptop discussed in Sec. III-A, the proposed algorithm took less than 10 seconds to search for the target pattern approximated by 41 basis functions in the 756×756 grey-scale image.

V. CONCLUSION

In this paper, we have obtained the optimal steerable basis for representing all rotated versions of a given function $h(x, y)$. We name this optimal representation “Fourier-Argand” representation. With this representation, rotating a 2D function amounts to multiplying the basis functions by pure complex exponentials. This simple rotation-covariance enables efficient computation of the convolution between a 2D signal and all rotated versions of h . With all its useful properties, the Fourier-Argand representation brings in the possibility to accurately approximate highly direction-selective

filters as well as complicated patterns. We further devised an algorithm to estimate the local orientation of the pattern in an image.² The estimation is very accurate even if the image suffers from large noise. As an example, we calculated, analytically, the Fourier-Argand representation of an elongated Gaussian function, and demonstrated its use for detection of ridge. Besides, we showed the relation between the Fourier-Argand representation and the Radon transform that leads to an efficient algorithm for computing the Fourier-Argand representation of a general *digital* pattern. We demonstrated the use of this algorithm by applying it to the pattern recognition problem, where the pattern is allowed to have an arbitrary orientation.

Another possible application of the Fourier-Argand representation is the design of rotation-covariant (complex) wavelets [19]. For example, Unser *et al.* generalised the Riesz transform (multi-dimensional extension of the Hilbert transform) to higher-orders in [28], which is used to construct steerable wavelets. The optimality of the Fourier-Argand representation suggests that it is a more compact representation of objects independent of their orientation. For example, in many vision tasks, the low level filters contain rotated versions of the same filter, which is heavily redundant [29]. The Fourier-Argand representation could, in principle, be useful for eliminating this redundancy. The rotation-covariance of the Fourier-Argand representation suggests that it could also be useful for designing rotation-invariant features [30].

²<http://www.ee.cuhk.edu.hk/~tblu/demos/FourierArgand/http://www.ee.cuhk.edu.hk/~tblu/demos/FourierArgand/>

APPENDIX A
PROOF OF THEOREM 1

Since the approximation is an orthogonal projection of the target function h onto the approximation space, we can rewrite the square of the approximation error as follows

$$e_N^2 = \|h\|^2 - \sum_{n=1}^N \frac{1}{2\pi} \int_{-\pi}^{\pi} |\langle {}^a h, \varphi_n \rangle|^2 d\alpha. \quad (33)$$

From the rotation-covariance property (9) of the Fourier-Argand representation (7), we have

$$\begin{aligned} & \frac{1}{2\pi} \int_{-\pi}^{\pi} |\langle {}^a h, \varphi_n \rangle|^2 d\alpha \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \left| \sum_k \langle H_k, \varphi_n \rangle e^{-ik\alpha} \right|^2 d\alpha \\ &= \sum_k \sum_{k'} \langle H_k, \varphi_n \rangle \langle \varphi_n, H_{k'} \rangle \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{i(k'-k)\alpha} d\alpha \\ &= \sum_k |\langle \varphi_n, H_k \rangle|^2. \end{aligned} \quad (34)$$

Substituting (34) back into (33), we find that minimising the approximation error amounts to the following maximisation problem under the orthonormality condition,

$$\begin{aligned} & \max_{\varphi_1, \varphi_2, \dots, \varphi_N} \sum_{n=1}^N \sum_k |\langle \varphi_n, H_k \rangle|^2 \\ & \text{s.t.} : \langle \varphi_n, \varphi_{n'} \rangle = \delta_{n-n'}. \end{aligned} \quad (35)$$

We define $\lambda_k = \langle H_k, H_k \rangle \geq 0$. Note that $\lambda_{-k} = \lambda_k$ if h is real-valued. We choose $k_1, k_2, \dots, k_N \in \mathbb{Z}$ such that

$$\lambda_k \leq \lambda_{k_n}, \text{ for all } n \in \{1, 2, \dots, N\} \text{ and } k \notin \{k_1, k_2, \dots, k_N\},$$

then the solution to (35) is given by $\varphi_n = H_{k_n} / \sqrt{\lambda_{k_n}}$, and in which case, the minimum approximation error is $e_N^2 = \|h\|^2 - \sum_{n=1}^N \lambda_{k_n}$.

We prove this by induction on N . First, if $N = 1$, there is

$$\begin{aligned} \sum_k |\langle \varphi_1, H_k \rangle|^2 &= \sum_k \lambda_k \left\langle \varphi_1, \frac{H_k}{\sqrt{\lambda_k}} \right\rangle^2 \\ &\leq \max_k (\lambda_k) \langle \varphi_1, \varphi_1 \rangle = \max_k (\lambda_k), \end{aligned}$$

where we used the orthonormality $\langle \frac{H_k}{\sqrt{\lambda_k}}, \frac{H_{k'}}{\sqrt{\lambda_{k'}}} \rangle = \delta_{k-k'}$. The inequality becomes an equality iff $k_1 = \arg \max_k \lambda_k$ and $\varphi_1 = H_{k_1} / \sqrt{\lambda_{k_1}}$. Next, suppose the theorem holds for $N - 1$, then

$$\begin{aligned} & \sum_{n=1}^N \sum_k |\langle \varphi_n, H_k \rangle|^2 \\ &= \sum_{n=1}^{N-1} \sum_k |\langle \varphi_n, H_k \rangle|^2 + \sum_k |\langle \varphi_N, H_k \rangle|^2 \\ &\leq \sum_{n=1}^{N-1} \lambda_{k_n} + \sum_{k \notin \{k_1, k_2, \dots, k_{N-1}\}} \lambda_k \left\langle \varphi_N, \frac{H_k}{\sqrt{\lambda_k}} \right\rangle^2 \\ &\leq \sum_{n=1}^{N-1} \lambda_{k_n} + \max_{k \notin \{k_1, k_2, \dots, k_{N-1}\}} \lambda_k, \end{aligned}$$

where we used the same trick on the second term as we did for $N = 1$. The equality is reached iff $k_N = \arg \max_{k \notin \{k_1, k_2, \dots, k_{N-1}\}} \lambda_k$ and $\varphi_N = H_{k_N} / \sqrt{\lambda_{k_N}}$. This completes the proof.

ACKNOWLEDGEMENT

The authors would like to thank the reviewers and the associate editor for their suggestions which improved the paper significantly.

REFERENCES

- [1] J. Canny, "A computational approach to edge detection," in *Readings in Computer Vision*. Amsterdam, The Netherlands: Elsevier, 1987, pp. 184–203.
- [2] S. Sarkar and K. L. Boyer, "On optimal infinite impulse response edge detection filters," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 13, no. 11, pp. 1154–1171, Nov. 1991.
- [3] W. T. Freeman and E. H. Adelson, "The design and use of steerable filters," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 13, no. 9, pp. 891–906, Sep. 1991.
- [4] M. Jacob and M. Unser, "Design of steerable filters for feature detection using Canny-like criteria," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 26, no. 8, pp. 1007–1019, Aug. 2004.
- [5] E. P. Simoncelli and H. Farid, "Steerable wedge filters for local orientation analysis," *IEEE Trans. Image Process.*, vol. 5, no. 9, pp. 1377–1382, Sep. 1996.
- [6] P. Arbeláez, M. Maire, C. Fowlkes, and J. Malik, "Contour detection and hierarchical image segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 5, pp. 898–916, May 2011.
- [7] J. Staal, M. D. Abramoff, M. Niemeijer, M. A. Viergever, and B. van Ginneken, "Ridge-based vessel segmentation in color images of the retina," *IEEE Trans. Med. Imag.*, vol. 23, no. 4, pp. 501–509, Apr. 2004.
- [8] A. C. Bovik, M. Clark, and W. S. Geisler, "Multichannel texture analysis using localized spatial filters," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 12, no. 1, pp. 55–73, Jan. 1990.
- [9] D. G. Lowe *et al.*, "Object recognition from local scale-invariant features," in *Proc. 7th IEEE Int. Conf. Comput. Vis.*, Sep. 1999, vol. 99, no. 2, pp. 1150–1157.
- [10] F. Ullah and S. Kaneko, "Using orientation codes for rotation-invariant template matching," *Pattern Recognit.*, vol. 37, no. 2, pp. 201–209, Feb. 2004.
- [11] G. Turin, "An introduction to matched filters," *IEEE Trans. Inf. Theory*, vol. It-6, no. 3, pp. 311–329, Jun. 1960.
- [12] Y.-N. Hsu and H. H. Arsenault, "Optical pattern recognition using circular harmonic expansion," *Appl. Opt.*, vol. 21, no. 22, pp. 4016–4019, Nov. 1982.
- [13] Y.-N. Hsu and H. H. Arsenault, "Pattern discrimination by multiple circular harmonic components," *Appl. Opt.*, vol. 23, no. 6, pp. 841–844, Mar. 1984.
- [14] Q. Wang, O. Ronneberger, and H. Burkhardt, "Rotational invariance based on Fourier analysis in polar and spherical coordinates," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 31, no. 9, pp. 1715–1722, Sep. 2009.
- [15] Y. Sheng and H. H. Arsenault, "Method for determining expansion centers and predicting sidelobe levels for circular-harmonic filters," *J. Opt. Soc. Amer. A, Opt. Image Sci.*, vol. 4, no. 9, p. 1793, Sep. 1987.
- [16] A. Khotanzad and Y. H. Hong, "Invariant image recognition by Zernike moments," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 12, no. 5, pp. 489–497, May 1990.
- [17] T. Zhao and T. Blu, "Detecting curves in very noisy images using Fourier-Argand moments," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Sep. 2019, pp. 3011–3015.
- [18] J. R. Argand, *Essai Sur Une Manière De Représenter Les Quantités Imaginaires Dans Les Constructions Géométriques*. Paris, France: Gauthier-Villars, 1874.

- [19] B. Forster, T. Blu, D. Van De Ville, and M. Unser, "Shift-invariant spaces from rotation-covariant functions," *Appl. Comput. Harmon. Anal.*, vol. 25, no. 2, pp. 240–265, Sep. 2008.
- [20] J. Radon, "On the determination of functions from their integral values along certain manifolds," *IEEE Trans. Med. Imag.*, vol. MI-5, no. 4, pp. 170–176, Dec. 1986.
- [21] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. Comput. Vis.*, vol. 60, no. 2, pp. 91–110, Nov. 2004.
- [22] E. Ricci and R. Perfetti, "Retinal blood vessel segmentation using line operators and support vector classification," *IEEE Trans. Med. Imag.*, vol. 26, no. 10, pp. 1357–1365, Oct. 2007.
- [23] F. Zhao, Q. Huang, and W. Gao, "Image matching by normalized cross-correlation," in *Proc. Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, May 2006, vol. 2, p. 2.
- [24] A. D. Hoover, V. Kouznetsova, and M. Goldbaum, "Locating blood vessels in retinal images by piecewise threshold probing of a matched filter response," *IEEE Trans. Med. Imag.*, vol. 19, no. 3, pp. 203–210, Mar. 2000.
- [25] F. Luisier and T. Blu, "SURE-LET multichannel image denoising: Interscale orthonormal wavelet thresholding," *IEEE Trans. Image Process.*, vol. 17, no. 4, pp. 482–492, Apr. 2008.
- [26] F. Xue, F. Luisier, and T. Blu, "Multi-Wiener SURE-LET deconvolution," *IEEE Trans. Image Process.*, vol. 22, no. 5, pp. 1954–1968, May 2013.
- [27] J. Li, F. Luisier, and T. Blu, "PURE-LET image deconvolution," *IEEE Trans. Image Process.*, vol. 27, no. 1, pp. 92–105, Jan. 2018.
- [28] M. Unser and D. van de Ville, "Wavelet steerability and the higher-order Riesz transform," *IEEE Trans. Image Process.*, vol. 19, no. 3, pp. 636–652, Mar. 2010.
- [29] A. B. Chan and N. Vasconcelos, "Counting people with low-level features and Bayesian regression," *IEEE Trans. Image Process.*, vol. 21, no. 4, pp. 2160–2177, Apr. 2012.
- [30] T. Ojala, M. Pietikäinen, and T. Mäenpää, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 7, pp. 971–987, Jul. 2002.



Tianle Zhao (Student Member, IEEE) received the bachelor's degree in information engineering (TRC) from Shanghai Jiao Tong University, China, in 2013. He is currently pursuing the Ph.D. degree with the Department of Electronic Engineering, The Chinese University of Hong Kong, Hong Kong.

His research interests include steerable representations, multidimensional wavelets, and pattern matching.



Thierry Blu (Fellow, IEEE) was born in Orléans, France, in 1964. He received the "Diplôme d'ingénieur" from École Polytechnique, France, in 1986, and from Télécom Paris (ENST), France, in 1988, and the Ph.D. degree in electrical engineering from ENST in 1996, for a study on iterated rational filterbanks, applied to wideband audio coding.

He was with Biomedical Imaging Group, Swiss Federal Institute of Technology, Lausanne, Switzerland, from 1998 and 2007. He is currently

a Professor with the Department of Electronic Engineering, The Chinese University of Hong Kong. His research interests include wavelets, approximation and sampling theory, sparse representations, image denoising, biomedical imaging, optics, and wave propagation.

Dr. Blu has been a member of the IEEE Signal Processing Theory and Methods Technical Committee from 2008 to 2013. He is currently on the board of the *EURASIP Journal on Image and Video Processing*, *SIAM Journal on Imaging Sciences*, and a member of the IEEE Bio Imaging and Signal Processing Technical Committee. He was a recipient of two best paper awards from the IEEE Signal Processing Society in 2003 and 2006, respectively. He is also the coauthor of a paper that received a Young Author Best Paper Award (2009) from the same society. He was an Associate Editor for the IEEE TRANSACTIONS ON IMAGE PROCESSING from 2002 to 2006, the IEEE TRANSACTIONS ON SIGNAL PROCESSING from 2006 to 2010, and *Elsevier Signal Processing* from 2008 to 2011.