sciendo

JAISCR, 2021, Vol. 11, No. 1, pp. 69 – 80
10.2478/jaiscr-2021-0005

A NOVEL METHOD FOR INVARIANT IMAGE RECONSTRUCTION

Mirosław Pawlak[1,2*], Gurmukh Singh Panesar[1] and Marcin Korytkowski [3]

[1]*Department of Electrical and Computer Engineering, University of Manitoba, Canada*

[2]*Information Technology Institute, University of Social Sciences, Lodz, Poland*

[3]*Institute of Computational Intelligence, Czestochowa University of Technology, Czestochowa, Poland*

[*]*E-mail: Miroslaw.Pawlak@umanitoba.ca*

Submitted: 8th July 2020; Accepted: 5th October 2020

Abstract

In this paper we propose a novel method for invariant image reconstruction with the properly selected degree of symmetry. We make use of Zernike radial moments to represent an image due to their invariance properties to isometry transformations and the ability to uniquely represent the salient features of the image. The regularized ridge regression estimation strategy under symmetry constraints for estimating Zernike moments is proposed. This extended regularization problem allows us to enforces the bilateral symmetry in the reconstructed object. This is achieved by the proper choice of two regularization parameters controlling the level of reconstruction accuracy and the acceptable degree of symmetry. As a byproduct of our studies we propose an algorithm for estimating an angle of the symmetry axis which in turn is used to determine the possible asymmetry present in the image. The proposed image recovery under the symmetry constraints model is tested in a number of experiments involving image reconstruction and symmetry estimation.

Keywords: object representation, invariant features, symmetry, radial orthogonal moments, continuous symmetry, ridge regression

1 Introduction

It is commonly believed that shape defines the most important feature we perceive objects. In fact, it is the most discriminative property allowing to infer and classify real-world objects. In the computational shape theory we view the concept of shape as the property that can be described by measurable geometric and topological features such as area, position, orientation, distances between points, angles between lines, Euler number just to name a few [1, 2]. Based on such features the fundamental problem is to define the shape equivalence involving understanding the conditions under which two distinct objects have the same shape. This question is equivalent to finding a class of transformations that when applied to an object they do not change its shape. In the case of 2D patterns this class of transformations includes:

1 Translation - the position of the object has changed.

2 Rotation - the orientation of the object has changed.

3 Reflection - the object has been folded over the line of reflection.

4 Gliding - the composite transformation consisting of a reflection over a line followed by translation along that line.

It is clear that the above transformations define isometries, i.e., they preserve the shape of the object. Formally, we say that a transformation $\mathbf{T}$ is an isometry if it preserves the distance between two points. Then, two objects $\mathbb{O}_1, \mathbb{O}_2$ are said to be congruent when there is an isometric transformation $\mathbf{T}$ such that $\mathbf{T}(\mathbb{O}_1) = \mathbb{O}_2$. It is known [1, 3, 4] that the four above listed transformations define a complete set of possible isometries for planar objects.
The concept of transformations that preserve shape is closely related to the property of object symmetry. In fact, symmetry is the object characteristic that is defined by the invariance with respect to a given isometry. Thus, we say that an object $\mathbb{O}$ is symmetric with respect to the isometry $\mathbf{T}$ if $\mathbf{T}(\mathbb{O}) = \mathbb{O}$. Hence, $\mathbb{O}$ is invariant with respect to $\mathbf{T}$ if applying the isometry $\mathbf{T}$ leaves the object unchanged. Specializing this definition to the reflection (bilateral) isometry with respect a line $\mathbf{l}$ (denoted as $\mathbf{T}_{\mathbf{l}}^{\text{refl}}$) we say that an object has reflection symmetry when there is a line $\mathbf{l}$ such that $\mathbf{T}_{\mathbf{l}}^{\text{refl}}(\mathbb{O}) = \mathbb{O}$. Analogous types of symmetries can be defined with respect to other isometries. Furthermore, the aforementioned isometries can be combined to define more complex symmetries.
The symmetry property results in the congruential class of objects that share the same set of symmetries. For instance, the letters $\{B, C, D, E\}$ share the horizontal line reflection symmetry.

Figure 1. Objects with the perfect reflection symmetry and the unique line of symmetry.

In this paper we focus on the reflection symmetry being the most common symmetry type in nature and applications. Nevertheless, the presented results can be extended to other types of symmetries. Figure 1 depicts objects with the ideal reflection symmetry and the uniquely defined line of symmetry.

Natural objects rarely exhibit the perfect symmetry due to various measurement errors and the inherent imperfection of the original object. Hence, symmetry should be consider as a continuous feature rather than a binary one, i.e., one should assign a degree of symmetry to a given object rather to classify it to either symmetric or non-symmetric patterns. In Figure 2 we show a group of natural objects that reveal a certain degree of reflection symmetry. The degree may vary from one object to another confirming the statistical nature of the degree of symmetry.
Quantitative analysis of the degree of asymmetry, present in an object, is directly related to the measure of the difference between the given object and its symmetric counterpart. If the space of underlying objects has the inner product structure (Hilbert space) then the orthogonal projection of the object $\mathbb{O}$ onto the subspace of symmetric objects with respect to the isometry $\mathbf{T}$ is given by $\frac{\mathbb{O}+\mathbf{T}(\mathbb{O})}{2}$ [5, 6]. Then, it is clear that the distance between $\mathbb{O}$ and its closest symmetric counterpart is

$$\left|\left|\frac{\mathbb{O}-\mathbf{T}(\mathbb{O})}{2}\right|\right|, \tag{1}$$

where $||\cdot||$ denotes the metric in the assumed space of objects. The distance in (1) may serve (once it is properly normalized) as the measure of the degree of the object symmetry/asymmetry with respect to a fixed isometry $\mathbf{T}$.

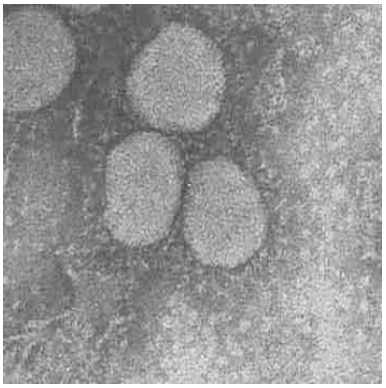

Figure 2. The corona virus objects revealing some degree of bilateral symmetry.

In this paper, we propose a novel method for invariant image reconstruction with the tunable degree of symmetry. This is achieved by combining the modern regression analysis with the theory of image invariants.

The rest of the paper is organized as follows. In Section 2 we examine the concept of reflection symmetry and the corresponding object representation in terms of invariant radial moments. In particular, the class of Zernike moments is introduced. Section 3 defines the image observation model and examines the computing aspects of Zernike moments for digital images. Section 4 provides an overview of regression analysis pertinent to the theory moment invariants used in this paper. Section 5 introduces our symmetry constrained reconstruction algorithm being the penalized version of the ridge regression estimate. Also an estimate of the axis of symmetry is proposed. Several experimental studies are presented to verify our methodology. In Section 6 we summarize our results and give further extensions.

2 Symmetry and Object Representation

Planar objects examined in this paper are assumed to be identified with the bivariate function $f(x,y)$ that represents the grey-level at the location (x,y). The function is defined on a certain subset $\mathbf{D}$ of $\mathbf{R}^2$. Then, the object is mirror (reflection/bilateral) symmetric if there is an axis of symmetry that divides the image into two identical reflected images. Formally, the image reveals the reflection symmetry if it belongs to the following class

$$\mathcal{S}^{\text{refl}} = \{f : f(x,y) = \mathbf{T}^{\text{refl}}_{\theta} f(x,y)\}, \qquad (2)$$

for some $\theta \in [0,\pi)$, where

$$\mathbf{T}^{\text{refl}}_{\theta} f(x,y) = f(x\cos(2\theta)+y\sin(2\theta), x\sin(2\theta)-y\cos(2\theta))$$

is the reflection of the image $f(x,y)$ with respect to the line of symmetry defined by the angle $\theta \in [0,\pi)$. Throughout the paper it is assumed, without loss of generality, that the coordinate system is located at the center of mass of the examined object.
It is also convenient to express the above symmetry constrain in terms of polar coordinates denoted as (ρ,φ). Hence, let $f(\rho,\varphi)$ be the version of the image function in polar coordinates. Then, the reflection symmetry transform in (2) reads as

$$\mathbf{T}^{\text{refl}}_{\theta} f(\rho,\varphi) = f(\rho, 2\theta-\varphi). \qquad (3)$$

The corresponding symmetry requirement for the image function becomes $f(\rho,\varphi) = f(\rho,2\theta-\varphi)$ for some angle θ. Common mirror symmetry classes are: vertical symmetry ($\theta = \pi/2$), horizontal symmetry ($\theta = 0$), and diagonal symmetry ($\theta = \pi/4$).
In many problems related to object recognition and description it is of interest to find the shape of an object from indirect measurements that define the so-called shape descriptors [2]. Efficient shape descriptors should posses some required properties such as:

- Completeness: the descriptors should represent uniquely the object information content.
- Invariance: the descriptors should be invariant for some or all isometry transformations.
- Robustness: the descriptors should be resistant to various object degradation processes like noise and sampling.

Moments and functions of moments define shape descriptors that posses some of the aforementioned useful properties [7, 8, 2]. The classical geometric moments of the so-called complex form are defined as

$$C_{pq}(f) = \iint_{\mathbf{D}} z^p z^{*q} f(x,y)\,dxdy,$$

where (p,q) is the moment order and $z = x+jy$, $z^* = x-jy$. Throughout the paper it is assumed that the object domain $\mathbf{D}$ is the unit disk. Then, the polar coordinates form of $C_{pq}(f)$ is given by

$$C_{pq}(f) = \int_0^1 \int_0^{2\pi} \rho^{p+q} e^{j(p-q)\varphi} f(\rho,\varphi)\rho\,d\rho\,d\varphi. \quad (4)$$

This form of $C_{pq}(f)$ and (3) allow us to obtain the complex moment of the image reflected by the axis tilted by the angle θ. Hence, we have

$$C_{pq}(\mathbf{T}^{\text{refl}}_{\theta} f) = e^{j(p-q)2\theta} C^*_{pq}(f).$$

Therefore, the mirror symmetry property implies that the following holds

$$e^{j(p-q)2\theta} C^*_{pq}(f) = C_{pq}(f). \qquad (5)$$

Since the analogous property holds for the rotation transformation therefore the magnitudes $|C_{pq}(f)|$ of

$C_{pq}(f)$ can be used as invariant features with respect to all rotations and reflections. Nevertheless, the symmetry property in (5) puts some restrictions on the moment order (p,q) such that moments of certain orders are not able to discriminate between different symmetric objects [7]. Moreover, complex moments are not orthogonal and as such they do not satisfy the completeness property. Also, they lack of robustness, i.e., inability to uniquely recover the image function from noisy observations. There is also little known about the accuracy of computing $C_{pq}(f)$ from digital and noisy data.

These shortcomings of the complex moments can be overcome by the extending the definition in (4) to moments with respect some general class of base functions. Hence, let

$$Z_{pq}(f) = \int_0^1 \int_0^{2\pi} V_{pq}^*(\rho,\varphi) f(\rho,\varphi)\rho d\rho d\varphi, \quad (6)$$

be a general class of moments with respect to a class of functions $\{V_{pq}(\rho,\varphi)\}$, where $V_{pq}^*(\rho,\varphi)$ is the complex conjugate of $V_{pq}(\rho,\varphi)$ defined on the unit disk, i.e., for $0 \le \rho \le 1$, $0 \le \varphi \le 2\pi$.

Particularly attractive is the choice of $\{V_{pq}(\rho,\varphi)\}$ that forms a class of orthogonal and complete functions, i.e., when

$$\int_0^1 \int_0^{2\pi} V_{pq}^*(\rho,\varphi) V_{p'q'}(\rho,\varphi)\rho d\rho d\varphi = \lambda_{pq}\delta_{pp'}\delta_{qq'}, \quad (7)$$

for some constant λ_{pq} and where $\delta_{ij} = 1$ if $i = j$ and 0 otherwise.

There is a large number of orthogonal sets defined over the unit disk. We can derive an unique set by putting some symmetry conditions on $V_{pq}(\rho,\varphi)$ and additional restrictions on the indices (p,q). It was proved in [9] that a basis that is invariant for any rotation of axes must be of the form

$$V_{pq}(x,y) = R_p(\rho)e^{jq\varphi}, \quad (8)$$

where $R_p(\rho)$ is a radial orthogonal polynomial of degree p and q defines the angular order. There are various ways of selecting $R_p(\rho)$ and important examples are Fourier-Mellin, pseudo-Zernike and Zernike radial bases [7, 8]. Among the possible choices for $R_p(\rho)$ there is only one orthogonal set, the set of Zernike functions, for which $R_p(\rho) = R_{pq}(\rho)$ is the radial orthogonal polynomial of degree $p \ge |q|$ such that $p - |q|$ is even [9]. Hence, the Zernike polynomial $R_{pq}(\rho)$ has no powers of ρ lower than $|q|$. The integers p, q are the degree and the azimuth order of the polynomial $R_{pq}(\rho)$, respectively. Also note that the Zernike functions satisfy (7) with $\lambda_{pq} = \pi/(p+1)$. The explicit form of the Zernike radial polynomial $R_{pq}(\rho)$ can be given as it has surprising relationship to the classical Jacobi polynomials [8]. Hence, it is known that

$$R_{pq}(\rho) = \rho^{|q|} P^{(0,|q|)}_{(p-|q|)/2}(2\rho^2 - 1),$$

where $P_k^{(a,b)}(t)$ is the $k-$degree orthogonal Jacobi polynomial on $[-1,1]$ with the parameters $a,b > -1$. This identity leads to some useful symmetry properties of the radial Zernike polynomial $R_{pq}(\rho)$, i.e., $R_{p,-q}(\rho) = R_{pq}(\rho)$ and $R_{pq}(-\rho) = (-1)^q R_{pq}(\rho)$. The figure below depicts several Zernike functions $V_{pq}(x,y)$ for various values of p,q.

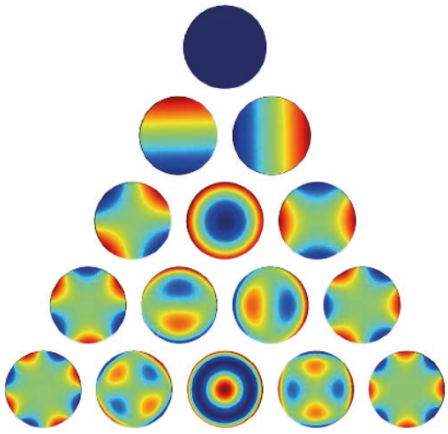

Figure 3. The collection of Zernike functions for various degrees p and the repetition index q.

Thus, the Zernike moment of order (p,q) is defined as

$$Z_{pq}(f) = \frac{p+1}{\pi} \int_0^{2\pi} \int_0^1 f(\rho,\varphi) e^{-jq\varphi} R_{pq}(\rho)\, \rho\, d\rho d\varphi, \quad (9)$$

where the normalization factor was included in the definition.

The fundamental property of Zernike moments is the easiness to impose the basic symmetry transformations into the radial moment formula in (9). In fact, applying the symmetry mappings in (3) to (9) one obtains the following relationship

$$Z_{pq}(\mathbf{T}_\theta^{\text{refl}} f) = e^{-2jq\theta} Z_{pq}^*(f). \quad (10)$$

The symmetry conditions can be now easily expressed in terms of the Zernike moment $Z_{pq}(f)$.

Hence, the image f exhibits the reflection symmetry with respect to the line of angle $\theta \in [0,\pi)$ if

$$e^{-2jq\theta} Z^*_{pq}(f) = Z_{pq}(f). \tag{11}$$

The above formula puts some constrains on the admissible set of Zernike moments. Hence, (11) holds if $Z_{pq}(f) \neq 0$ and moreover, we have

$$\arg(Z_{pq}(f)) = -q\theta. \tag{12}$$

In particular, this implies that if the horizontal mirror symmetry holds then $Z_{pq}(f)$ is real. The important difference between these restrictions and the ones established for complex moments is that the symmetry invariance conditions of radial moments influence only the angular order q defining the Zernike moments $\{Z_{pq}(f)\}$. Furthermore, the orthogonality property of Zernike functions allows us to group moments such that we can uniquely characterize the symmetry condition in (11) and in the same time to recover the image. In conclusion, the Zernike moments posses the aforementioned essential properties of completeness and invariance. This allows us to write the unique representation of the image function as the T-term expansion in terms of $\{V_{pq}(x,y)\}$

$$f_T(x,y) = \sum_{p=0}^{T} \sum_{q=-p}^{p} Z_{pq}(f) V_{pq}(x,y), \tag{13}$$

where the summation is taken with respect to the admissible pairs (p,q), i.e., $p \geq |q|$ and $p-|q|$ is even. Moreover, $V_{pq}(x,y)$ can be expressed in polar coordinates as in (8). It is also useful to express the representation in (13) in the real-valued form. A simple algebra employing the aforementioned properties of $V_{pq}(x,y)$ gives the following equivalent representation

$$f_T(\rho,\varphi) =$$

$$\sum_{p=0}^{T} \sum_{q=0}^{p} \left(A_{pq}(f)\cos(q\varphi) + B_{pq}(f)\sin(q\varphi)\right) R_{pq}(\rho), \tag{14}$$

where $A_{pq}(f), B_{pq}(f)$ are the Zernike moments corresponding to the cos and sin parts of the complex exponential, respectively. Also $p-q$ = even and $p \geq q \geq 0$.

In the explicit form we have

$$\begin{aligned}
A_{p0}(f) &= \frac{1}{\pi} \int_0^{2\pi} \int_0^1 f(\rho,\varphi) R_{p0}(\rho) \rho d\rho d\varphi \\
A_{pq}(f) &= \frac{2(p+1)}{\pi} \int_0^{2\pi} \int_0^1 f(\rho,\varphi)\cos(q\varphi) R_{pq}(\rho) \rho d\rho d\varphi, \\
B_{p0}(f) &= 0 \\
B_{pq}(f) &= \frac{2(p+1)}{\pi} \int_0^{2\pi} \int_0^1 f(\rho,\varphi)\sin(q\varphi) R_{pq}(\rho) \rho d\rho d\varphi.
\end{aligned} \tag{15}$$

The formulas in (13), (2) and (15) can be used to obtain the equivalent $T-$term representation for the reflected image with respect to the symmetry line of the angle θ. We will denote this representation as $f^{\theta}_T(\rho,\varphi)$. Recalling (13) and the formula in (10) we obtain

$$f^{\theta}_T(\rho,\varphi) = \sum_{p=0}^{T} \sum_{q=-p}^{p} e^{-2jq\theta} Z^*_{pq}(f) V_{pq}(\rho,\varphi). \tag{16}$$

On the other hand the reflected counterpart of (15) takes the form

$$f^{\theta}_T(\rho,\varphi) = \sum_{p=0}^{T} \sum_{q=0}^{p} \left(A^{\theta}_{pq}(f)\cos(q\varphi) + B^{\theta}_{pq}(f)\sin(q\varphi)\right) R_{pq}(\rho), \tag{17}$$

where

$$\begin{aligned}
A^{\theta}_{pq}(f) &= A_{pq}(f)\cos(q\pi)\cos(2q\theta) + B_{pq}(f)\cos(q\pi)\sin(2q\theta) \\
B^{\theta}_{pq}(f) &= A_{pq}(f)\cos(q\pi)\sin(2q\theta) - B_{pq}(f)\cos(q\pi)\cos(2q\theta).
\end{aligned} \tag{18}$$

are the real-valued Zernike moments of the reflected image.

It is worth noting that the L_2 distance between two images f,g can be easily expressed in term of Zernike moments. In fact, by virtue of the orthogonality and Parseval's formula we can approximate the L_2 distance $\|f-g\|^2$ as

$$\|f-g\|^2 \approx \sum_{p=0}^{T} \sum_{q=-p}^{p} |Z_{pq}(f) - Z_{pq}(g)|^2, \tag{19}$$

where $Z_{pq}(f)$ and $Z_{pq}(g)$ are Zernike moments of f and g, respectively.

This combined with (10) allows us to obtain the distance between the image function and its reflected version in terms of the Zernike moments, i.e., we have

$$\|f - \mathbf{T}^{\text{refl}}_{\theta} f\|^2 \approx \sum_{p=0}^{T} \sum_{q=-p}^{p} |Z_{pq}(f) - e^{-2jq\theta} Z^*_{pq}(f)|^2. \tag{20}$$

Furthermore, applying this to (1) would lead to the Zernike moments based distance between the given object f and its closest symmetric counterpart (with respect to the angle θ). In fact, the distance is just $\|f-\mathbf{T}_\theta^{\text{refl}} f\|/2$ and this due to (20) can be expressed in terms of the Zernike moments. Furthermore, the proper normalization of $\|f-\mathbf{T}_\theta^{\text{refl}} f\|/2$, i.e.,

$$\|f-\mathbf{T}_\theta^{\text{refl}} f\|/2\|f\|$$

yields the symmetry index $\mathbf{SI}(\theta)\in[0,1]$ such that $\mathbf{SI}(\theta)=0$ if the object f is perfectly reflection symmetric with respect to the symmetry line of the angle θ.

3 Image Degradation and Zernike Moments Computing

In Section 2 we have assumed the full knowledge of the image function f defined on the unit disk. In real-world applications one does not have a complete information about the image function f and must verify the question whether f reveals some type of symmetry observing only its discrete and distorted version. Hence, we observe the distorted and digital version of the image function over the pixel set $\{x_k,y_l;1\le k,l\le n\}$ according the following model

$$Y_{kl}=\mathbb{D}f(x_k,y_l), \tag{21}$$

for $1\le k,l\le n$, where $\mathbb{D}f(x,y)$ is a certain deformation operator. This may include noise, blurring and missing data just to name a few. The linear blurring and additive noise is an example of such deformation

$$Y_{kl}=\iint_{\mathbf{D}}\Psi(x_k-x,y_l-y)f(x,y)dxdy+\varepsilon_{kl},$$

where $\Psi(x,y)$ is the point-spread function of the given imaging system and ε_{kl} is the noise process. This image degradation model plays important role in confocal microscopy and medical imaging [7, 6, 10]. In this paper we mostly focus on the diskretization error represented by the n^2 pixels set that defines the digital image resolution. Hence, the data in (21) are observed on the square grid of the edge width Δ, i.e., $x_k-x_{k-1}=y_l-y_{l-1}=\Delta$ and (x_k,y_l) is the center of the pixel (k,l). Note that Δ is of order $1/n$ and the image size is $N=n^2$. In Figure 4 we show the employed pixel configuration within the circular image support. This is often referred to as the inner circle digitization [8, 11].

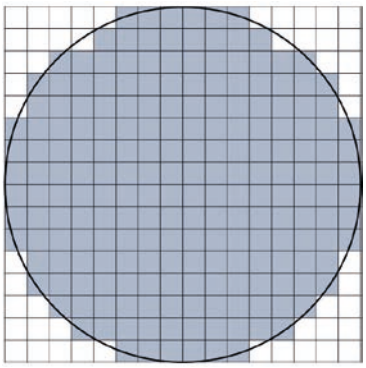

Figure 4. Inner circle digitization of the circular domain.

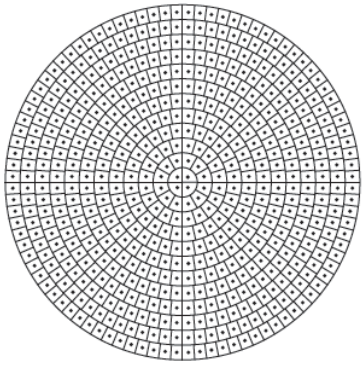

Figure 5. The mapping from (x,y) coordinates to (ρ,φ) coordinates.

For the given data set in (21) the fundamental problem is to evaluate the Zernike moment $Z_{pq}(f)$ defined in (9). The numerical integration method takes the following generic form

$$\widetilde{Z}_{pq}=\frac{p+1}{\pi}\sum_{k=1}^{n}\sum_{l=1}^{n}Y_{kl}\iint_{p_{kl}}V_{pq}^*(\rho,\varphi)\rho\,d\rho d\varphi, \tag{22}$$

where p_{kl} is the pixel centred at (x_k,y_l) and we used the modified notation $\widetilde{Z}_{pq}$ instead of $\widetilde{Z}_{pq}(f)$. The simplest way of evaluating the integral in (22) is to apply the first order integration scheme. For the higher-order integration methods we refer to [8, 12]. Thus, applying the first order numerical integration, we have the following approximation

$$\iint_{p_{kl}}V_{pq}^*(\rho,\varphi)\rho\,d\rho d\varphi\approx V_{pq}^*(\rho_k,\varphi_l)\rho_k\Delta\rho_k\Delta\varphi_l,$$

where $\Delta\rho_k=\rho_k-\rho_{k-1}$, $\Delta\varphi_l=\varphi_l-\varphi_{l-1}$. It is worth noting that the polar coordinates (ρ_k,φ_l) should be

obtained from the proper $(x,y) \mapsto (\rho,\varphi)$ mapping that generates the polar tiling of the unit disk. The figure below shows an example of such mapping, see [8, 11, 12] for more details.

The estimate in (22) has the straightforward analog for the real-valued Zernike moments defined in (15). The aforementioned estimates of $\{Z_{pq}(f)\}$ can be directly used to define the object reconstruction from the distorted data. Hence, the following estimate of the image function results from the combination of (22) and (13)

$$\widetilde{f}_T(x,y) = \sum_{p=0}^{T} \sum_{q=-p}^{p} \widetilde{Z}_{pq} V_{pq}(x,y). \tag{23}$$

The analogous estimate can be defined for the real-valued reconstruction in (2).

It was shown [8, 11, 12] that under the additive noise model the estimate $\widetilde{f}_T(x,y)$ can converge (in the mean L_2 norm sense) to the true image function with the rate $O(\Delta^{2/3})$ given that the truncation parameter T is of order $\Delta^{-2/3}$, where Δ is the pixel width. Thus, if the image resolution increases ($\Delta \to 0$) one can recover the true image with the aforementioned rate.

The above reconstruction method presents the traditional way of recovering images from the orthogonal moments as it is based on the direct estimation of the Zernike moments. This method was extensively examined in [8, 11, 12].

In this paper we propose the alternative approach that is utilizing modern regression algorithms like ridge regression. We will show that this leads to a more accurate reconstruction method. Moreover, the regression strategy allows us to incorporate shape constraints such as symmetry. This is discussed in Sections 4 and 5.

4 Regression Analysis

Our strategy to obtain new estimates of Zernike moments is based on the extension of the standard least squares regression analysis. To do so let us re-write the reconstruction formula in (2) in a slightly simplified form

$$f_T(\rho,\varphi) = \sum_{p=0}^{T} \sum_{q=0}^{p} \left(A_{pq}\cos(q\varphi) + B_{pq}\sin(q\varphi)\right) R_{pq}(\rho), \tag{24}$$

where we used the simpler notation A_{pq}, B_{pq} instead of $A_{pq}(f), B_{pq}(f)$. Let us assume the observation model

$$Y_{kl} = f(x_k, y_l) + \varepsilon_{kl}, \tag{25}$$

for $1 \le k,l \le n$ and $\{\varepsilon_{kl}\}$ is the zero-mean noise process. Furthermore, let us represent the pixel locations $\{(x_k,y_l)\}$ in terms of polar coordinates $\{(\rho_k,\varphi_l)\}$, see Section 3 for the discussion of this issue.

Then, the ordinary least squares method to estimate A_{pq}, B_{pq} is obtained by the following minimization

$$\sum_{k,l=1}^{n} \left(Y_{kl} - \sum_{p=0}^{T} \sum_{q=0}^{p} \left(A_{pq}\cos(q\varphi_l) + B_{pq}\sin(q\varphi_l)\right) R_{pq}(\rho_k) \right)^2 \tag{26}$$

with respect to $\{A_{pq}, B_{pq}\}$. Here $\sum_{k,l=1}^{n}$ is the double summation with respect to all $1 \le k,l \le n$.

It will be useful in our future developments to rewrite the above formulas in the matrix form by introducing the vector notation. Hence, let

$$\beta = [A_{00}, A_{11}, \ldots, A_{TT}, B_{00}, B_{11}, \ldots, B_{TT}]^t$$

where t denotes the transpose. The vector β consists of the all unknown coefficients $\{A_{pq}, B_{pq}; p \le T\}$ that satisfy the restriction $p \ge q \ge 0$ and $p-q$ is even. Note that the vector β is $d-$dimensional with $d = (T+1)(T+2)/2$. Furthermore, let us denote $V^r_{pq}(\rho,\varphi) = R_{pq}(\rho)\cos(q\varphi)$ and $V^i_{pq}(\rho,\varphi) = R_{pq}(\rho)\sin(q\varphi)$. These are the real and imaginary parts of the Zernike functions, respectively. Then, one can define the following $N \times d$ matrix

$$X = [x_{11}, x_{12}, \ldots, x_{NN}]^t$$

where $N = n^2$ and

$$\begin{aligned} x_{kl} \;=\; & [V^r_{00}(\rho_k,\varphi_l), V^r_{11}(\rho_k,\varphi_l), \ldots, V^r_{TT}(\rho_k,\varphi_l), \\ & V^i_{00}(\rho_k,\varphi_l), V^i_{11}(\rho_k,\varphi_l), \ldots, V^i_{TT}(\rho_k,\varphi_l)] \end{aligned}$$

is the $d-$dimensional row vector. The $N \times d$ matrix X is often referred to as the design matrix.

Finally the observed data are summarized by the following $N-$dimensional vector

$$Y = [Y_{11}, \ldots, Y_{1n}, Y_{21}, ..., Y_{2n}, \ldots, Y_{nn}]^t.$$

With the above notation the formula in (26) can be written as the classical least squares criterion

$$L(\beta) = ||Y - X\beta||_2^2, \tag{27}$$

where $||\cdot||_2$ is the Euclidean distance. The minimum of $L(\beta)$ is given by

$$\widehat{\beta}_L = (X^t X)^{-1} X^t Y. \tag{28}$$

This forms the basic estimate of the Zernike coefficients. It is known that if the noise process $\{\varepsilon_{kl}\}$ is spatially uncorrelated and equal variance σ^2 for each position (k,l) then the estimate $\widehat{\beta}_L$ is unbiased with the variance given by

$$\mathrm{Var}[\widehat{\beta}_L] = \sigma^2 (X^t X)^{-1}. \tag{29}$$

The generalization of the estimate $\widehat{\beta}_L$ to its penalized version that reveals the numerical stability and sparsity will be discussed in Section 5. Clearly, one could use the estimate $\widehat{\beta}_L$ in the representation in (24) and to obtain the alternative reconstruction formula compared to the standard method in (23). The regression strategy has been rarely used in the image processing literature.

5 Penalized Regression Object Reconstruction and Symmetry Constraints

In this chapter we extend the classical regression analysis applied to the Zernike moments based image reconstruction to a statistical regularization method known as ridge regression. Next we examine how this modified regression strategy can incorporate symmetry constraints.

5.1 Ridge Regression Image Reconstruction from Zernike Moments

Thus far we have examined the case of multiple linear regression in which we assume that design matrix X is of the full rank d so that $(X^t X)^{-1}$ exists. In our study, however, we are composing X using covariates obtained from Zernike functions up to the order T yielding the matrix of the size $N \times d$, where d is of order T^2. These covariates are collinear thus making unfeasible to use the ordinary least squares for estimating Zernike moments. The regularized version of the least squares criterion in (27) is a general framework to impose specific restrictions on the sought solution. Hence, in order to avoid the singularity problem of $(X^t X)^{-1}$ one can consider the following penalized version of (27)

$$LR(\beta) = ||Y - X\beta||_2^2 + \lambda ||\beta||_2^2, \tag{30}$$

where $\lambda \geq 0$ is a regularization parameter that controls the impact of the penalty term. The minimization of $LR(\beta)$ yields the so-called ridge regression estimate of β [13]. Figure 6 illustrates the geometry of the ridge estimate where the design matrix has substantial negative correlation.

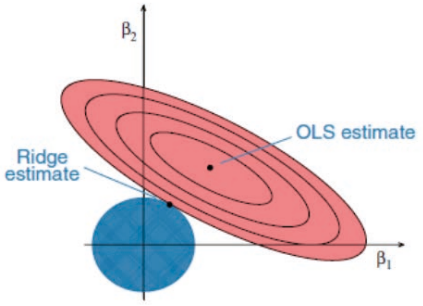

Figure 6. Ridge regression estimate compared with the ordinary least squares solution.

The explicit formula for the minimum of $LR(\beta)$ in (30) is given by

$$\widehat{\beta}_R = (X^t X + \lambda I)^{-1} X^t Y, \tag{31}$$

where I is the $d \times d$ identity matrix. The form of $\widehat{\beta}_R$ explains the regularization character of the ridge regression solution. In fact, for $\lambda \to 0$, the solution is approaching the standard least squares estimate $\widehat{\beta}_L$ in (28). On the other hand, for larger λ, the matrix $X^t X + \lambda I$ is well defined and therefore invertible regardless whether $X^t X$ has rank smaller than d. It is also useful to determine the bias and variance of the ridge estimate. First, let us observe that

$$\widehat{\beta}_R = (X^t X + \lambda I)^{-1} X^t X \widehat{\beta}_L. \tag{32}$$

Then, by virtue of the aforementioned properties of $\widehat{\beta}_L$ one can obtain

$$\mathbb{E}[\widehat{\beta}_R] = (X^t X + \lambda I)^{-1} X^t X \beta \tag{33}$$

and

$$\mathrm{Var}[\widehat{\beta}_R] = \sigma^2 A X^t X A^{-1}, \tag{34}$$

where $A = (X^t X + \lambda I)^{-1}$. The formula in (33) explains the shrinkage property of the ridge regression estimate. In fact, due to the orthogonality property in (7) the matrix $X^t X$ is approximately diagonal. In fact, it can be proved that

$$\sum_{(x_k, y_l) \in D} V^*_{pq}(x_k, y_l) V^*_{p'q'}(x_k, y_l) \Delta^2$$

is equal to

$$\frac{\pi}{p+1} \lambda_{pq} \delta_{pp'} \delta_{qq'} + O(\Delta^{1+\delta})$$

for some $0 < \delta < 1$. This represents the discrete orthogonality property being the counterpart of (7). By this and by the comparison of (28) with (31) we approximately have

$$\widehat{\beta}_R = \frac{1}{1+\lambda}\widehat{\beta}_L. \tag{35}$$

Since $\lambda > 0$, therefore the ridge regression estimate of the Zernike moments would shrink towards zero all small moment values as compared to the standard least squares solution. Furthermore, by the same reasoning as in (35) and the identity in (34) we can conclude that the estimate $\widehat{\beta}_R$ has smaller variability than the estimate $\widehat{\beta}_L$. In the case of natural images one is dealing with high-dimensional data and also the collinearity of the design matrix X is common. As a result, the variance is a dominating term of the mean square error of the estimate $\widehat{\beta}_R$ and this indicates that the regularization is necessary.

The ridge parameter λ can be selected by the $V-$fold cross validation and this provides the proper variance/bias tradeoff.

If X^tX is approximately diagonal, then by using (33) and (34) it is straightforward to write the formula for the mean square error of the estimate $\widehat{\beta}_R$. In fact, we have the following bias/variance decomposition

$$\mathbb{E}\left[||\widehat{\beta}_R - \beta||_2^2\right] = \frac{\lambda^2}{(1+\lambda)^2}||\beta||_2^2 + \frac{d}{(1+\lambda)^2}\sigma^2. \tag{36}$$

We note that the first term (representing the estimate bias) is increasing with λ, whereas the second term (representing the estimate variance) is decreasing with λ. Minimization of (36) leads to the explicit form for the optimal λ, i.e., we have

$$\lambda^* = \frac{d}{||\beta||_2^2}\sigma^2. \tag{37}$$

The true value of β in the above expression can be replaced by some pilot estimate, e.g., the ridge regression estimate with $\lambda = 1$. On the other hand the noise variance σ^2 can be evaluated by a class of universal techniques utilizing the difference of observations [10, 6]. This would lead to the explicit and practical estimate of λ.

All the aforementioned examination produces the efficient ridge regression estimate $\widehat{\beta}_R$ of the Zernike moments defined in (24). We will denote the resulting estimate of the image function as $\widehat{f}_T(x,y)$. It is informative to compare this reconstruction algorithm with the standard numerical integration based method introduced in (23). Figure 7 shows the reconstruction of the Mandril image for increasing values of the truncation parameter T, where the image resolution is 100×100. Similar experiments were conducted for a larger class of natural images revealing the great advantage of the ridge regression reconstruction method over the traditional one. This is summarized in Figure 8 where the reconstruction error versus the maximal moment order T is plotted for the assumed class of images.

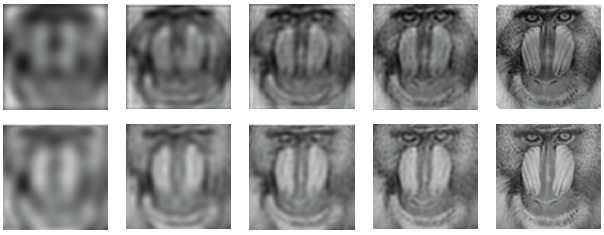

Figure 7. Image reconstruction based on the standard method $\widetilde{f}_T(x,y)$ (top panel) in (23) and the ridge regression method $\widehat{f}_T(x,y)$ (lower panel).

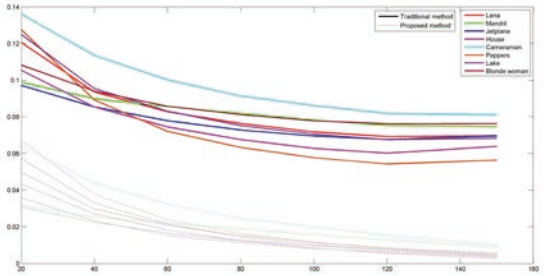

Figure 8. Reconstruction error for the standard reconstruction method $\widetilde{f}_T(x,y)$ and the ridge regression estimate $\widehat{f}_T(x,y)$ versus T for a large class of natural images.

5.2 Image Reconstruction with Enforced Symmetry

In this Section we wish to further extend the penalized regression analysis to the form that allows to incorporate the symmetry constrain. Hence, for the fixed axis of reflection symmetry parametrized by the angle θ let us define the following risk function

$$LS(\beta) = ||Y - X\beta||_2^2 + \lambda||\beta||_2^2 + \lambda_s||\beta - \beta(\theta)||_2^2, \tag{38}$$

where $\beta(\theta)$ is the version of β corresponding to the reflected image with respect to the line of symme-

try of the angle θ. Furthermore, $\lambda_s \geq 0$ is the regularization parameter that is controlling the degree of symmetry we wish to accept in the reconstructed image. Clearly, if $\lambda_s = 0$ then we recover the ridge regression risk function in (30).

The formula for the transformed $\beta(\theta)$ results from equations given in (18). This can be written in the matrix form as

$$\beta(\theta) = T(\theta)\beta, \tag{39}$$

where $T(\theta)$ is the $d \times d$ matrix having the block-wise structure. Owing to (18) the (p,q) block is of the form

$$\begin{bmatrix} \cos(q\pi)\cos(2q\theta) & \cos(q\pi)\sin(2q\theta) \\ \cos(q\pi)\sin(2q\theta) & -\cos(q\pi)\cos(2q\theta) \end{bmatrix}.$$

The direct minimization of $LS(\beta)$ yields the following double regularized estimate of the Zernike moments

$$\widehat{\beta}_{RS} = \left(X^t X + \lambda I + \lambda_s (I - T(\theta))^t (I - T(\theta))\right)^{-1} X^t Y. \tag{40}$$

Clearly, for $\lambda_s = 0$ we obtain the ridge regression estimate $\widehat{\beta}_R$. As we have already observed the ridge regression tuning parameter λ plays the important role in reducing the variance of the ridge regression estimate. On the other hand, the symmetry regularization parameter λ_s controls the shape property of the reconstructed image in terms of its degree of bilateral symmetry. Hence, larger λ_s implies the more symmetric solution. The symmetry tuning parameter λ_s can be set by the user or selected based on a preliminary symmetry testing procedure. In the latter case one should perform a formal testing and use the result of the test to choose the proper value of λ_s. Hence, if the symmetry hypothesis is accepted then λ_s should be small, otherwise the larger value of λ_s is recommended. In this case the power of the test should indicate the value of λ_s. Further details of the hypothesis testing driven approach for selecting λ_s will be studied elsewhere. We refer to [10] for the theory of testing for image symmetries.

The estimate $\widehat{\beta}_{RS}$ in (42) used in (2) yields to the image reconstruction formula $\widehat{f}_T(x,y;\lambda_s)$ parametrized by λ_s, where we assume that the ridge parameter λ was already specified. As we have already discussed $\widehat{f}_T(x,y;0)$ gives the ridge regression estimate $\widehat{f}_T(x,y)$ that is not symmetry regularized. The estimate $\widehat{f}_T(x,y;\lambda_s)$ with larger λ_s leads to a nested sequence of image estimates with the increasingly degree of symmetry.

To verify the usefulness of the reconstruction method $\widehat{f}_T(x,y;\lambda_s)$ we choose a class of images with varying levels of visual asymmetry as is is shown in Figure 9. For instance, the butterfly image is almost perfectly symmetric while the other images reveal only the partial symmetry. The ellipse shape with the symmetry line lying outside the image is entirely non-symmetric.

Figure 9. Testing images with various degrees of bilateral symmetry.

The simulation results are obtained for the following range of $\lambda_s \in \{100, 200, 300, 400, 500\}$. The Zernike moment order T was set to 100. Figure 10 shows reconstructed images as a result of enforcing bilateral symmetry as well as the effect of the increasing values of λ_s. The line of symmetry, i.e., the parameter θ was selected based on the method examined in the next section.

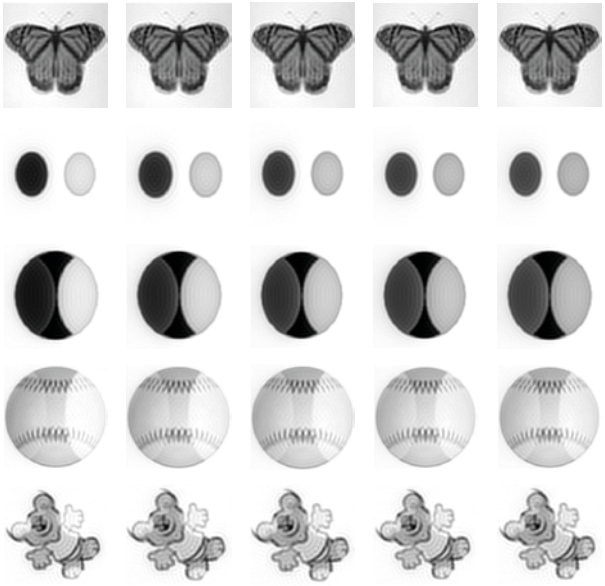

Figure 10. The symmetry regularized Zernike moments reconstruction method $\widehat{f}_T(x,y;\lambda_s)$ with the increasing values of λ_s.

Figure 11 depicts the square error (versus λ_s) between reconstructed image using $\widehat{f}_T(x,y;\lambda_s)$ and its reflected version obtained with respect to the se-

lected symmetry axis. This represents the amount of asymmetry present in the given image. The plotted curves tend to flatten as λ_s increases beyond a certain value. Clearly, the fastest decrease of the asymmetry curves is observed for the butterfly image as being the most symmetric. The bear image in Figure 9 is an example of the image with missing data. It is seen that the symmetry driven regularization is able to recover the missing information.

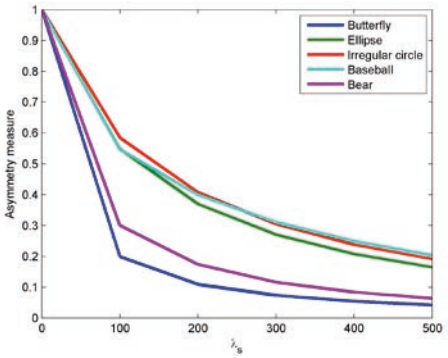

Figure 11. Symmetrization index.

5.3 Symmetry Axis Estimation

Thus far, the choice of the angle θ of reflection symmetry was set arbitrary or left to the user that can specify θ based on some a priori information about the examined class of objects. In this Section we wish to give the automatic, data-driven choice of θ. In order to find the proper θ we can use some measure of the disagreement between moments of the given image and its reflected version with respect to a given value of θ. A natural such measure is defined in (20) which due to Parseval's formula can be expressed in terms of Zernike moments. This leads to the following symmetry distance

$$C(\theta) = ||\beta - \beta(\theta)||_2^2. \tag{41}$$

This is clearly equal to the symmetry penalty term in (38), where $\beta(\theta)$ is the version of β corresponding to the reflected image with respect to the line parametrized by the angle θ.

Since we have estimated the Zernike moments β by the ridge regression method therefore we can replace β in (41) by the ridge regression estimate $\widehat{\beta}_R$. The ridge regression tuning parameter λ can be specified in the way as it was discussed in the previous subsection, see (37).

Since by virtue of (39) we have $\beta(\theta) = T(\theta)\beta$ therefore the angle of the symmetry axis can be estimated by minimizing the following empirical counterpart of $C(\theta)$, i.e., we have

$$\widehat{\theta} = \arg\min_{\theta} \widehat{C}(\theta), \tag{42}$$

where $\widehat{C}(\theta) = ||\widehat{\beta}_R - T(\theta)\widehat{\beta}_R||_2^2$.

The above minimization problem is one dimensional and can be easily solved by either plotting $\widehat{C}(\theta)$ or by applying some efficient minimization algorithm for single variable functions such as the golden section search method.

Figure 12 depicts the results of using the above approach for symmetry estimation. Images with different values of the angle of symmetry have been applied. The lower panel in Figure 12 plots $\widehat{C}(\theta)$ versus θ revealing local and global minima of the symmetry distance function. The point of the global minima yields our estimate $\widehat{\theta}$.

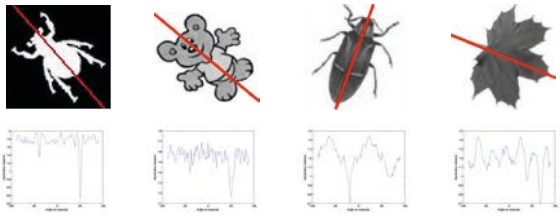

Figure 12. Symmetry line estimation.

The statistical precision of the estimate $\widehat{\theta}$ in (42) would be an important problem for future research. In [6] the theory of estimating of θ was established with the use of the standard method of estimating Zernike moments, see (22). Based on these results we can conjecture that the estimate $\widehat{\theta}$ in (42) has some favourable statistical properties. Hence, we can expect that it reveals the optimal parametric rate, i.e.,

$$\widehat{\theta} = \theta^* + O_P(\Delta),$$

where Δ is the edge width of the digital image with the resolution $n \times n$, i.e., when Δ is of order $1/n$. Here θ^* denotes the true unknown angle of reflection symmetry. Also, the symbol $O_P(\cdot)$ denotes the convergence in probability. This is the challenging semi-parametric estimation problem as the image function $f(x,y)$ is unknown and has the nonparametric nature.

6 Concluding Remarks and Extensions

In this paper we presented the unified penalized regression framework for object reconstruction with the imposed degree of symmetry. This strategy has many potential applications in object recognition, compression and understanding. It is plain that many questions, both theoretical and implementation type, remain to be addressed. In fact, we have examined the object symmetry within a single object. It would be interesting to extend this methodology to the matching symmetry being symmetry between two corresponding objects. Furthermore, the proposed methodology could be extended to composite symmetries and symmetries for 3D objects. In the latter case a number of symmetry classes is much larger than in the 2D setting.

References

[1] U. Grenander, Elements of Pattern Theory. John Hopkins University Press, 1996.

[2] D. Zhang and G. Lu, Review of shape representation and description techniques, Pattern Recognition, vol. 37, pp. 1–19, 2004.

[3] H. Weyl, Symmetry. Princeton University Press, 1952.

[4] J. Rosen, Symmetry in Science: An Introduction to the General Theory. Springer-Verlag, 1995.

[5] H. Zabrodsky, S. Peleg, and D. Avnir, Symmetry as a continuous feature, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 17, pp. 1154–1166, 1995.

[6] N. Bissantz, H. Holzmann, and M. Pawlak, Improving PSF calibration in confocal microscopic imaging - estimating and exploiting bilateral symmetry, The Annals of Applied Statistics, vol. 4, pp. 1871–1891, 2010.

[7] J. Flusser, T. Suk, and B. Zitova, 2D and 3D Image Analysis by Moments. John Wiley and Sons, 2017.

[8] M. Pawlak, Image Analysis by Moments. Publishing House of Technical University of Wroclaw, 2006.

[9] A. B. Bhatia and E. Wolf, On the circle polynomials of Zernike and related orthogonal sets, Proceedings of Cambridge Philosophical Society, vol. 50, pp. 40–48, 1954.

[10] N. Bissantz, H. Holzmann, and M. Pawlak, Testing for image symmetries-with application to confocal microscopy, IEEE Transactions on Information Theory, vol. 55, pp. 1841–1855, 2009.

[11] Y. Xin, M. Pawlak, and S. Liao, Accurate computation of Zernike moments in polar coordinates, IEEE Transactions on Image Processing, vol. 16, pp. 581–587, 2007.

[12] M. Pawlak and S. Liao, On the recovery of a function on a circular domain, IEEE Transactions on Information Theory, vol. 48, pp. 2736–2753, 2002.

[13] G. James, D. Witten, T. Hastie, and R. Tibshirani, An Introduction to Statistical Learning. Springer, 2013.

Miroslaw Pawlak received the Ph.D. and D.Sc. degrees in computer engineering from Wrocław University of Technology, Wrocław, Poland. He is currently a Professor at the Department of Electrical and Computer Engineering, University of Manitoba, Winnipeg, MB, Canada. He has held a number of visiting positions in North American, Australian, and European Universities. He was at the University of Ulm, University in Goettingen and Marburg University as an Alexander von Humboldt Foundation Fellow. His research interests include statistical signal processing, machine learning, and nonparametric modeling. Among his publications in these areas are the books Image Analysis by Moments (Wrocław Univ. Technol. Press, 2006), and Nonparametric System Identification (Cambridge Univ. Press, 2008), coauthored with Prof. Włodzimierz Greblicki. Dr. Pawlak has been an Associate Editor of the Journal of Pattern Recognition and Applications, Pattern Recognition, International Journal on Sampling Theory in Signal and Image Processing, Journal of Artificial Intelligence and Soft Computing Research, Opuscula Mathematica and Statistics in Transition-New Series.

Gurmukh Singh Panesar received the M.Tech in Electronics and Instrumentation Engineering from Sri Guru Granth Sahib World University, Punjab, India and pursuing the M.Sc. in Electric and Computer Engineering from the University of Manitoba, Winnipeg, MB, Canada. He published a paper on tumour diagnosis using the method of moment. His current fields of interest are symmetry estimation, regression analysis, and machine learning.

Marcin Korytkowski received the Ph.D. and degree in computer science from Częstochowa University of Technology, Poland, in 2007. Currently he is an associate professor at Częstochowa University of Technology. He has published over 70 technical papers. His present research interests include deep learning architectures and their applications in computer security, databases and image retrieval. He is involved in organization of various scientific events and conferences.