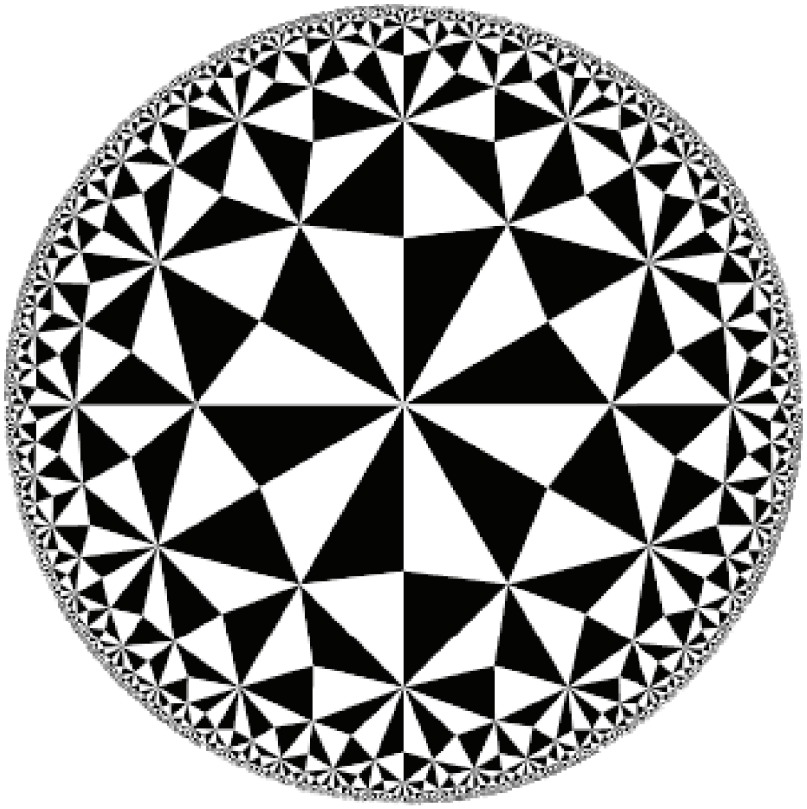
Mirosław Pawlak

Image Analysis by Moments: Reconstruction and Computational Aspects



Mirosław Pawlak

Image Analysis by Moments: Reconstruction and Computational Aspects

OFICYNA WYDAWNICZA POLITECHNIKI WROCŁAWSKIEJ
WROCŁAW 2006

Recenzent
Dariusz Uciński

Opracowanie redakcyjne i korekta
Hanna Jurek

Skład komputerowy
Miroslaw Pawlak

Projekt okładki
Miroslaw Pawlak

OFICYNA WYDAWNICZA POLITECHNIKI WROCŁAWSKIEJ
50-370 Wrocław, Wybrzeże Wyspiańskiego 27

ISBN 83–7085–913–5

Drukarnia Oficyny Wydawniczej Politechniki Wrocławskiej. Zam. 2006.

Contents

Preface

Image processing has become a critical component in contemporary science and technology and has numerous applications ranging from nanotechnologies, astronomy, medicine, vision psychology, remote sensing, security screening, and military industry to the digital communication technologies.

This book is about the problem of image characterization and analysis via the domain-based descriptors utilizing the mathematical theory of moments and orthogonal polynomial bases. The moments and orthogonal polynomials are classical concepts in mathematical analysis and statistics not only due to their simplicity and own elegance but also for the extraordinary range of subjects and applications where they have illuminated. In fact they may fulfill the well known Occam's Razor principle saying that "when you have two competing theories which make exactly the same predictions, the one that is simpler is the better."

There already exist a number of monographs in the general area of image analysis and in image feature selection and image descriptors in particular [63, 112, 160, 161, 203]. These books give excellent overview of the various concepts of image descriptor techniques and include good illustrations of applications. The aim of this text is to complement the existing research in several ways. Firstly, we focus on a specific class of image descriptors stemming from the theory of moments and orthogonal polynomials of two variables as well as the theory of invariance. Secondly, we give the fundamental accuracy analysis of the introduced moment descriptors. This includes assessing the performance of the descriptors with respect to common data deformations such as discretization and noise. We also examine an error due to the geometric nature of the image plane and this is referred to as the geometric error. This type of error is explained by finding the connection between the accuracy issue with the analytic number theory of lattice point approximations.

Several new techniques to increase the accuracy and efficiency of moment descriptors are also proposed. We utilize these results for solving the problem of reconstruction of noisy images from orthogonal moments. The theory developed reveals fundamental trade-offs between the aforementioned types of errors. This

leads to the issue of an intrinsic dimensionality of a feature vector which yields an optimal representation of an image.

The obtained fundamental results are employed to tackle two important problems of image analysis. The first one concerns symmetry detection. Detection of symmetries and symmetry-based representations constitute important practical issues that have not received much attention in the image analysis literature. We develop formal statistical procedures to test symmetries existing in an image observed in the presence of noise. In the second application we propose a novel watermarking system able to cope with geometric transformations. Watermarking algorithms form a critical part of modern multimedia systems where data hiding and copyright protection issues are very essential.

The present monograph relies entirely on the original results of the author and his collaborators. Here we list the salient features of this monograph:

- A detailed overview of the problem of invariant object descriptors is given with main emphasis on the issues of accurate numerical computation and noise sensitivity.
- Properties of a class of orthogonal moments stemming from the theory of classical orthogonal polynomials are thoroughly examined. This includes questions of numerical efficiency, reconstruction power from the computed moments, robustness to noise, and automatic selection of optimal number of moments
- An extension of the above results to the case of generalized moments and orthogonal moments calculated in the digital domain is given.
- Properties of radial orthogonal moments are examined. A thorough error analysis of the moments is conducted including numerical error, the accuracy in the presence of noise, and geometric error. In the latter case the error analysis is explained by relating the accuracy issue to the analytic number theory of lattice point approximations.
- The reconstruction power of radial moments is studied and optimal convergence rates for image recovery from moments are derived.
- Fast and high quality algorithms for computing radial moments are derived and empirically tested.
- Statistical methodology, utilizing the theory of radial moments, for testing image symmetry is developed. This includes testing image rotational and reflectional symmetries. The limit distributions for test statistics are established.
- A novel watermarking system, utilizing the theory of radial moments, is proposed. The system exhibits high robustness to geometric attacks, noise, filtering, and data compression. Watermark imbedding and extraction algorithms

are developed. Extensive empirical tests on the system accuracy are conducted.

The book is organized as follows. In Chapter 1 we give an introduction to basic issues related to image descriptors and moment descriptors especially. The heart of this book is Chapters 2 through 4, in which original results on the accuracy analysis of orthogonal moment descriptors are presented. In particular, Chapter 2 is devoted to the image analysis by orthogonal moments stemming from the theory of classical orthogonal polynomials. On the other hand, Chapters 3 and 4 are concerned with a class of radial polynomials of the invariant form. The importance and basic properties of this class of moment descriptors are discussed in Chapter 3, whereas Chapter 4 examines the reconstruction properties of the radial invariant moments. Chapter 5 gives a careful treatment of the problem of efficient and accurate computation of the radial invariant moments. Chapters 6 and 7 show that many of the basic ideas developed in Chapters 2 through 4 have immediate applications in two important problems of image processing. Hence in Chapter 6 we develop the statistical theory for testing an image symmetry. Chapter 7, in turn, is concerned with an application of radial moments to the problems of data hiding and watermarking. All of the aforementioned results are based on the original contributions and ideas of the author and his collaborators.

Over the years, my own work has benefited greatly from the advice and support of a number of friends and colleagues with interest in ideas of non-parametric estimation, pattern recognition and imaging. There are too many names to list here, but special mention is due to Prof. Włodzimierz Greblicki who has introduced me to these research areas. Without his help and encouragement my consciousness would be still today in a narrow, finite dimensional "parametric" world. Our collaboration has resulted in numerous joint papers on pattern recognition, non-parametric estimation, and nonlinear system identification.

I would also like to thank Prof. Ewaryst Rafajłowicz, who has been a great source of helpful advice and who has encouraged me to write this book. We have been involved in a number of joint research projects concerning smoothing techniques in signal sampling, non-parametric control charts, and pattern recognition.

Important contributions to the ideas addressed in this book have been made by a succession of research collaborators and students. This, in particular, includes Prof. Zygmunt Hasiewicz who has contributed greatly to my research papers in the area of nonlinear system identification and wavelet modeling. Furthermore, Prof. Uli Stadtmüller has been a great host during my several visits to the University of Ulm. We have developed a fruitful long standing research collaboration in the areas of non-parametric modeling for signal analysis and statistics.

I am deeply indebted to Prof. Zygmunt Hasiewicz and Prof. Adam Krzyżak for a careful review of this book. I am also grateful to Prof. Adam Krzyżak, Dr. Yongqing Xin, Prof. Simon Liao, Dr. Hajo Holzmann, Prof. Andrzej Kozek, and Prof. Pradeepa Yahampath for joint research papers and numerous discussions.

Finally, but by no means least, I would like to thank Mount-first Ng for helping me with a number of typesetting problems. I also thank Mrs. Elżbieta Żurawska-Łuczyńska for being a very supportive and patient editor.

Research presented in this monograph was partially supported by the Polish Foundation for Science and NSERC of Canada.

Chapter 1

Object Representation in Feature Spaces: Moment-Based Techniques

An overview of moment-based methods for invariant representation of the structure of complex objects, such as images, is presented. The aim of the invariant representation is to characterize an object and all its affine transformed versions by a point in a high-dimensional feature space obtained from Cartesian or orthogonal moments. These methods are derived from the general theory of moments and orthogonal polynomials which is widely used throughout mathematics, physics and statistics. A basic moment theory is reviewed with main emphasis on a class of orthogonal moments. We discuss both Cartesian (geometric) and orthogonal moments. The moment representation methods are discussed with respect to invariance, symmetry and sensitivity to noise and discretization. The fundamental problem of image reconstruction from moments is introduced.

1.1. Introduction

When analyzing complex objects as images, it is vital to have a simple, effective, robust, computationally efficient approach that can represent salient object features. One way of achieving this task is to manually select a number of important characteristics that are directly related to the objective of a study, as for instance face characteristics in the face recognition problem. Such a strategy is clearly application dependent and difficult to automatize. The feature extraction approach aims at representing an object as a point in a finite dimensional space (feature space) such that different views of the object correspond to the same point in

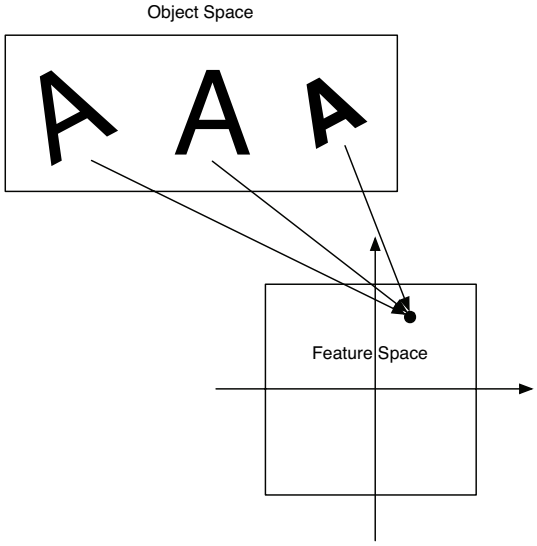


Figure 1.1: A feature space for 2D objects obtained by translation, scaling, and rotation of the original image

the feature space. In Fig. 1.1 we illustrate this concept by showing the original image and its two versions obtained through translation, scaling, and rotation.

The feature space representation of an image has several advantages. First, we obtain a significant reduction of dimensionality without loosing important information about the original object. If the representation is carefully chosen we can also obtain features which are relatively insensitive to noise and occlusion. In fact, a class of orthogonal moment descriptors, examined extensively in this book, can efficiently compress an original, e.g., $64 \times 64 = 4096$-dimensional image to several nearly independent components. Moments have been used to distinguish between shapes of different characters, aircrafts, chromosomes, and industrial parts [161], [160]. Furthermore, this type of image features provides a complete object representation that is invariant to similarity transformations. The principal issue, being thoroughly examined in this book, of the feature space representation is the capability of reconstruction of the original object from a finite and often noisy set of object descriptors. This is the inverse problem which is typically ill-posed and requires some regularization methods.

1.2. Geometric Moments

Simple geometric properties of an image such as area, position, and orientation can be easily computed from a set of linear functionals of the image called geometric moments. Hence let $f : \Omega \subset R^2 \rightarrow R$, Ω being some compact set, be an image function describing a real scene, such that $0 \leq f(x, y)$ represents an intensity of the image at a spatial position $(x, y) \in \Omega$, where Ω is often called the image plane.

We define the (p, q)-th moment of $f(x, y)$ as follows

$$m_{pq} = \iint_{\Omega} x^p y^q f(x, y) dx dy. \tag{1.1}$$

A set of moments up to order N consists of all moments m_{pq} such that $0 \leq p+q \leq N$ and if (p, q) are non-negative integers then the set contains $(N + 1)(N + 2)/2$ elements. The moments up to order two provide basic geometric characteristics of $f(x, y)$. In fact m_{00} defines the total mass of $f(x, y)$, whereas the moments (m_{00}, m_{10}, m_{01}) are sufficient to calculate the center of mass of the image $f(x, y)$, i.e.,

$$\overline{x} = \frac{m_{10}}{m_{00}}, \qquad \overline{y} = \frac{m_{01}}{m_{00}}. \tag{1.2}$$

Assuming that the image is somewhat elongated we can also find the orientation of the image defined as the axis of least inertia, i.e., the line for which the integral $\iint_{\Omega} r^2(x, y) f(x, y) dx dy$ is minimum, where $r(x, y)$ is the perpendicular distance from the point (x, y) to the line sought after. It can be easily shown that the axis of least inertia can be obtained from the second order moments (m_{20}, m_{11}, m_{02}) [199]. Higher order moments can provide additional detailed information about the image such as an orientation of the axis of least inertia and symmetry properties. The above geometric concepts have particularly simple interpretation for binary images, i.e., when $f(x, y) = \mathbf{1}_D(x, y)$, where D is a subset of Ω and $\mathbf{1}_D(x, y)$ is the indicator function of D.

It is also worth mentioning that moments can be easily calculated from projections of the image. In fact, it is known that the horizontal, vertical, and diagonal projections contain all the information needed to compute the moments $\{m_{pq}, 0 \leq p + q \leq 2\}$. These moments, in turn, provide all the information that we need to obtain the position and orientation of the planar image. This fact is fundamental to tomographic reconstruction in medical imaging [33, 157].

A central moment corresponding to m_{pq} is defined as follows

$$\mu_{pq} = \iint_{\Omega} (x - \overline{x})^p (y - \overline{y})^q f(x, y) dx dy. \tag{1.3}$$

The central moments are equivalent to the regular geometric moments of an image that has been shifted such that the image centroid $(\overline{x}, \overline{y})$ is at the origin. As a result, central moments are invariant to translations of the image. Throughout the book, without loss of generality, we will assume that the origin coincides with the centroid of the image, i.e., m_{pq} will be identified with μ_{pq}. The use of moments for image analysis and object recognition was initiated by Hu [101] and described in great detail in [160, 199, 203]. It is also worth mentioning that moments play an important role in the various branches of mathematics such as analysis, probability theory, and statistics, see [137] for a comprehensive overview of the moment theory from viewpoint of mathematics.

Moment Invariants

As we have already observed the central moments are invariant to the image translation. It is a fundamental problem in image analysis to find image descriptors which are invariant under general affine transformations of the image plane

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \alpha & a \\ b & \beta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} x_0 \\ y_0 \end{bmatrix}. \tag{1.4}$$

It is easy to show that the following normalized central moments

$$\nu_{pq} = \frac{\mu_{pq}}{\mu_{00}^{1+(p+q)/2}}, \quad p+q \geq 2 \tag{1.5}$$

are scale-invariant, i.e., invariant under transformation (1.4) with $a = b = 0$.

A significant step in finding moment invariants under transformation (1.4) was made by Hu [101], who employing the general theory of algebraic invariants [164] derived the following first seven invariant moments

$$\begin{aligned}
\psi_1 &= \mu_{20} + \mu_{02}, && (1.6)\\
\psi_2 &= (\mu_{20} - \mu_{02})^2 + 4\mu_{11}^2, && (1.7)\\
\psi_3 &= (\mu_{30} - 3\mu_{12})^2 + (3\mu_{21} - \mu_{03})^2, && (1.8)\\
\psi_4 &= (\mu_{30} + \mu_{12})^2 + (\mu_{21} + \mu_{03})^2, && (1.9)\\
\psi_5 &= (\mu_{30} - 3\mu_{12})(\mu_{30} + \mu_{12})[(\mu_{30} + \mu_{12})^2 - 3(\mu_{21} + \mu_{03})^2] \\
&\quad + (3\mu_{21} - \mu_{03})(\mu_{21} + \mu_{03})[3(\mu_{30} + \mu_{12})^2 - (\mu_{21} + \mu_{03})^2], && (1.10)\\
\psi_6 &= (\mu_{20} - \mu_{02})[(\mu_{30} + \mu_{12})^2 - (\mu_{21} + \mu_{03})^2] \\
&\quad + 4\mu_{11}(\mu_{30} + \mu_{12})(\mu_{21} + \mu_{03}), && (1.11)\\
\psi_7 &= (3\mu_{21} - \mu_{03})(\mu_{30} + \mu_{12})[(\mu_{30} + \mu_{12})^2 - 3(\mu_{21} + \mu_{03})^2] \\
&\quad - (\mu_{30} - 3\mu_{12})(\mu_{21} + \mu_{03})[3(\mu_{30} + \mu_{12})^2 - (\mu_{21} + \mu_{03})^2], && (1.12)
\end{aligned}$$

where we assumed that $\mu_{00} = 1$, i.e., that instead of the moments $\{\nu_{pq}\}$ we can equivalently use $\{\mu_{pq}\}$. The descriptors $\psi_1, \ldots, \psi_7$ are supposed to be invariant under transformation (1.4) but they are actually invariant under shifting, scaling and rotation transformations. This class of transformations does not include, e.g., a shear transformation. The shear transformation parallel to the x-axis is represented (as in (1.4)) by the following matrix

$$\begin{bmatrix} 1 & a \\ 0 & 1 \end{bmatrix}. \tag{1.13}$$

The corrected version of Hu's invariant moments was established in [202] (and rediscovered in [50]). The following set is composed of the first four affine transformation invariants

$$\begin{aligned}
I_1 &= (\mu_{20}\mu_{02} - \mu_{11}^2)/\mu_{00}^4, && (1.14)\\
I_2 &= (\mu_{30}^2\mu_{03}^2 - 6\mu_{30}\mu_{21}\mu_{12}\mu_{03} + 4\mu_{30}\mu_{12}^3 + 4\mu_{03}\mu_{21}^3 \\
&\quad - 3\mu_{21}^2\mu_{12}^2)/\mu_{00}^{10}, && (1.15)\\
I_3 &= (\mu_{20}\mu_{21}\mu_{03} - \mu_{20}\mu_{12}^2 - \mu_{11}\mu_{30}\mu_{03} + \mu_{11}\mu_{21}\mu_{12} + \mu_{02}\mu_{30}\mu_{12} \\
&\quad - \mu_{02}\mu_{21}^2)/\mu_{00}^7, && (1.16)\\
I_4 &= (\mu_{20}^3\mu_{03}^2 - 6\mu_{20}^2\mu_{11}\mu_{12}\mu_{03} - 6\mu_{20}^2\mu_{02}\mu_{21}\mu_{03} + 9\mu_{20}^2\mu_{02}\mu_{12}^2 \\
&\quad - 18\mu_{20}\mu_{11}\mu_{02}\mu_{21}\mu_{12} - 8\mu_{11}^3\mu_{30}\mu_{03} - 6\mu_{20}\mu_{02}^2\mu_{30}\mu_{12} \\
&\quad + 12\mu_{20}\mu_{11}^2\mu_{21}\mu_{03} + 6\mu_{20}\mu_{11}\mu_{02}\mu_{30}\mu_{03} + 9\mu_{20}\mu_{02}^2\mu_{21}^2 \\
&\quad + 12\mu_{11}^2\mu_{02}\mu_{30}\mu_{12} - 6\mu_{11}\mu_{02}^2\mu_{30}\mu_{21} + \mu_{02}^3\mu_{30}^2)/\mu_{00}^{11}. && (1.17)
\end{aligned}$$

It is worth noting that the above moment invariants are given as a highly nonlinear composition of the geometric moments up to order three. The invariants based on higher order moments have also been obtained, e.g., employing geometric moments up to the ninth-order we can generate 52 moment invariants [143]. In practice, however, we use (due to sensitivity to noise) lower order invariants. There have also been recent developments in defining invariant moments with respect to a linear blurring operation [48, 49, 51–53].

There are a number of limitations in the practical use of the aforementioned moment invariants. In fact the following is a list of their shortcomings.

- Due to the small number of invariants it is difficult to use them in representing complex objects. This is an important issue is some applications, e.g., in digital watermarking systems we require a large number of object descriptors in order to embed a robust watermark, see Section 7 for a discussion of this problem.

- The nonlinear nature of the invariants makes them potentially unstable. A small error in the computation of $\{\mu_{pq}\}$ may result in unpredictable errors of the invariants. This is a particularly serious issue when we observe the image in the presence of noise.
- The geometric moments are defined for analog images, and so are the invariants $\psi_1, ..., \psi_7$ and $I_1, ..., I_4$. For digital images, the invariance property is only approximately satisfied due to numerical error involved in moment computation [144].
- For images revealing a certain type of symmetry a number of moments are equal to zero. For instance, for images symmetric about the y-axis, i.e., when $f(x, y) = f(-x, y)$, $(x, y) \in \Omega = [-1, 1]^2$, we have $\mu_{11} = \mu_{12} = \mu_{30} = 0$. This leads to a great reduction in the number of non-zero invariants. In the extreme case of rotationally symmetric objects all of the seven Hu's invariants are zero.

1.3. Orthogonal Moments

The geometric moments are formed using a monomial basis set $\{x^p y^q\}$. This is a complete set of functions on Ω but not orthogonal. The lack of orthogonality of $\{x^p y^q\}$ results in high correlation between corresponding moments $\{m_{pq}\}$ yielding a highly redundant representation of the image. Since different powers of $x^p y^q$ differ very little from each other therefore the geometric moments $\{m_{pq}\}$ have a great difficulty in distinguishing between different patterns. Teague [219] proposed a simple idea of overcoming the above shortcomings of $\{m_{pq}\}$ by replacing $\{x^p y^q\}$ by a complete orthogonal polynomial system $\{V_{pq}(x, y)\}$. The orthogonality of $\{V_{pq}(x, y)\}$ means that we have

$$\iint_{\Omega} V_{pq}(x, y) V_{p'q'}(x, y) w(x, y) dx dy = 0$$

for all $(p, q) \neq (p', q')$, where $w(x, y)$ is the weight function.

There are several ways of constructing two-variable counterparts of orthogonal polynomials [126]. The simplest strategy is to take the direct product of two systems of orthogonal polynomials in one variable. Hence let $\{P_p(x)\}$ be the orthogonal polynomial basis with respect to the weight function $w_1(x)$ which is defined on the set Ω_1. Similarly, let $\{Q_q(y)\}$ be the orthogonal polynomial basis with respect to the weight function $w_2(y)$ which is defined on the set Ω_2 and let $\Omega = \Omega_1 \times \Omega_2$. Then it can be shown that

$$V_{pq}(x, y) = P_p(x) Q_q(y) \tag{1.18}$$

is the orthogonal polynomial of degree $p+q$ with respect to the weight function $w_1(x)w_2(y)$ which is defined on Ω. Hence $\{P_p(x)Q_q(y), p, q = 0, 1, \dots\}$ constitutes the orthogonal polynomial basis on Ω.

More general orthogonal polynomial bases with respect to the weight function $w(x, y)$ can be obtained [126] by applying the Gram-Schmidt orthogonalization process to the sequence of monomials

$$1, x, y, x^2, xy, y^2, x^3, x^2y, \dots.$$

The orthogonal polynomial of order $p+q$ is defined as the polynomial with the highest term x^py^q obtained in the orthogonalization process.

The following two important cases of orthogonal polynomial bases are examined thoroughly in this book.

1. *Classical orthogonal polynomials in two variables.* Here $\Omega = [-1, 1]^2$ and in (1.18) $\{P_p(x)\}$, $\{Q_q(y)\}$ are classical orthogonal polynomials like Legendre, Chebyshev, Gegenbauer, and Jacobi polynomials [206, 217]. The classical one-variable orthogonal polynomials share a number of fundamental properties. One of them is the basic three-term recurrence relation which must be satisfied by any classical one-variable orthogonal polynomial, i.e., we have

$$P_{k+1}(x) = (x - \alpha_k)P_k(x) - \beta_k P_{k-1}(x), \qquad k = 0, 1, \dots, \tag{1.19}$$

where $P_{-1}(x) = 0$, $P_0(x) = 1$ and α_k, β_k are constants. This formula yields efficient algorithms for generating the classical orthogonal polynomials and consequently designing two dimensional orthogonal polynomials according the prescription given in (1.18).

2. *Radial orthogonal functions.* Here $\Omega = \{(x, y) : x^2 + y^2 \leq 1\}$ is the unit disk and the general form of radial orthogonal functions is the following

$$V_{pq}(x, y) = R_{pq}(\rho)e^{jq\theta}, \tag{1.20}$$

where $\rho = \sqrt{x^2 + y^2}$, $\theta = \arctan(y/x)$, and $R_{pq}(\rho)$ is a polynomial in ρ. This particular form of orthogonal functions originates from the seminal work of Bhatia and Wolf [13], who gave a procedure for constructing polynomials which form complete sets on the unit disk and are invariant in form with respect to rotation of axes. It was proved in [13] that each such set must be of the form as in (1.20) with q being an integer positive, negative or zero and the polynomial $R_{pq}(\rho)$ is of degree p, containing no power of ρ lower than $|q|$; moreover $R_{pq}(\rho)$ is either an even or odd polynomial

depending on whether q is even or odd. The specific form of polynomial $R_{pq}(\rho)$ depends on the choice of the radial weight function $w(\rho)$ with respect to which the functions $V_{pq}(x, y)$ are orthogonal. See Section 3 for further discussion of this issue.

The above considerations give the following definition of orthogonal moments with respect to the orthogonal complete polynomial system $\{V_{pq}(x, y)\}$.

$$\lambda_{pq} = \iint_{\Omega} V_{pq}(x, y) f(x, y) w(x, y) dx dy. \tag{1.21}$$

The issue of invariance for orthogonal moments can be resolved in various ways. First, one can find a relationship between orthogonal moments and geometric moments. Then we can use the theory of invariants discussed in Section 1.2 in order to find the corresponding invariants in the orthogonal moments domain. A simpler solution for translation- and scale-invariance can be obtained by shifting and scaling the image prior to the computation of the orthogonal moments. As we have already pointed out in Section 1.2 in order to get such a normalization we need to evaluate the geometric moments $\{m_{00}, m_{10}, m_{01}\}$. An important issue of rotational invariance remains to be addressed for the case of the classical orthogonal polynomial based moments defined by (1.18) and (1.21). On the other hand the radial orthogonal function based moments defined by (1.20) and (1.21 possess the inherent rotational invariance property. This is a primary reason for successful applications of the radial orthogonal moments in image representation and classification.

1.4. Shape From Moments

In practical applications of the theory of object descriptors one can only determine a finite number of them and then believe that the obtained features uniquely represent an image. This is generally not the case and in particular the uniqueness does not hold for the geometric moments. In fact it is well known that there exists an infinite variety of functions whose first N geometric moments coincide and a unique reconstruction of the image function $f(x, y)$ is impossible. Nevertheless, various approximation procedures exist [112] which aim at constructing a specific image function $g_N(x, y)$, e.g., a polynomial of order N, such that

$$\iint_{\Omega} x^p y^q g_N(x, y) dx dy = m_{pq}, \qquad 0 \leq p + q \leq N,$$

where $\{m_{pq}\}$ are known moments of $f(x, y)$.

The problem of whether the geometric moments can uniquely represent a single-variable function $f(x)$ defined on a finite interval is a classical topic in the approximation theory and is called the Hausdorff moment problem [224]. Hence let $m_p = \int_0^1 x^p f(x)dx$, $p = 0, 1, \ldots$ be the moments of a nonnegative integrable function $f(x)$ defined on $[0, 1]$. Hausdorff proved [224] that the following

$$\partial^r m_p = \sum_{l=0}^{r} (-1)^l \binom{r}{l} m_{p+l} > 0, \qquad p, r = 0, 1, 2, \ldots \tag{1.22}$$

is the necessary and sufficient condition for a unique representation of a given function $f(x)$ by its moments and more importantly this condition assures that for a given sequence $\{m_p\}$ (satisfying (1.22)) there exists a unique nonnegative $f(x)$ integrable on $[0, 1]$ whose moments coincide with $\{m_p\}$. The sequence $\{m_p\}$ satisfying (1.22) is called completely monotonic.

The extension of the Hausdorff moment problem to the case of functions of two variables was made in [99] where condition (1.22) is replaced by its two-dimensional counterpart

$$\partial_1^r \partial_2^h m_{pq} > 0, \qquad r, h, p, q = 0, 1, 2, \ldots, \tag{1.23}$$

where the operator ∂_1^r is defined as in (1.22) and applies to the first subscript p, and the second operator ∂_2^h applies independently to the second subscript q.

The Hausdorff moment theorem asserts that from an infinite set of completely monotonic moments $\{m_{pq}\}$ we can uniquely find the image function $f(x, y)$. This inverse problem, however, is not well posed. In fact, the inverse of the mapping

$$f \longmapsto \{m_{pq}\}$$

is not continuous [218]. Hence replacing $\{m_{pq}\}$ by $\{m_{pq} + \varepsilon_{pq}\}$, where $\{\varepsilon_{pq}\}$ is an arbitrarily small positive sequence, may result in the image function $f_\varepsilon(x, y)$ which is arbitrarily far from $f(x, y)$. This ill-posedness of the inverse moment problem has resulted in a number of regularization type solutions [218].

This instability of the inverse moment problem can be partially circumvented if one restricts the class of image functions. Hence let us consider a class of binary images $f(x, y) = \mathbf{1}_G(x, y)$, where G is an L-sided simply connected polygonal region in the image plane (see Fig. 1.2) with the L vertices represented by the complex numbers $z_1, z_2, \ldots, z_L$.

The following interesting result concerning the aforementioned polygons was proved by Davies in [30].

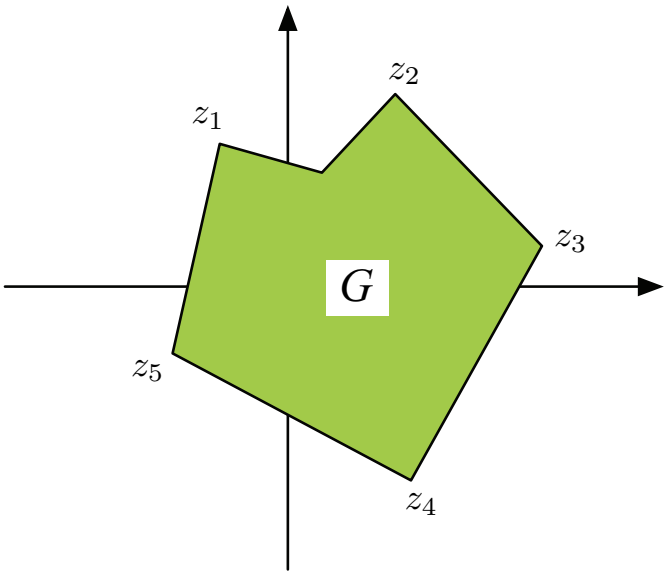


Figure 1.2: Planar object of the polygonal shape

Theorem 1.1. *Let $z_1, z_2, \ldots, z_L$ designate the vertices of an L-sided simply connected polygon G in the plane. Then we can find constants $a_1, a_2, \ldots, a_L$ depending upon $z_1, z_2, \ldots, z_L$ but independent of $h(z)$, such that for all $h(z)$ analytic in the closure of G,*

$$\iint_G h''(z)dxdy = \sum_{j=1}^{L} a_j h(z_j), \tag{1.24}$$

where $z = x + jy$ and $h''(z)$ is the second derivative of $h(z)$.

Using $h(z) = z^p$ in Theorem 1.1 we can readily obtain the following fundamental relationship

$$\tau_p = p(p-1)c_p = \sum_{j=1}^{L} a_j z_j^p, \tag{1.25}$$

where $c_p = \iint_G z^p dxdy$ is the p-th complex moment of the L-sided polygonal region G, and $\tau_0 = \tau_1 = 0$. It was shown in [157] that we can uniquely reconstruct G from the moments $c_0, c_1, \ldots, c_{2L-3}$ or equivalently from the numbers $\tau_0, \tau_1, \ldots, \tau_{2L-1}$. Various reconstruction algorithms based on the Davies theorem

have been reported in [62,65,86]. The issue of robustness to noise and discretization of the proposed algorithms remains to be addressed, see [61] for a recent study of this problem.

The above discussion on the shape recovery from the classical geometric moments reveals once again the serious limitations of this type of image descriptors. Having the choice of base functions for defining moments one can expect to be able to improve the reconstruction properties by employing the orthogonal moments $\{\lambda_{pq}\}$ introduced in (1.21). Indeed, the reconstruction strategy from orthogonal moments is very straightforward and does not need any regularization. Hence we simply represent an image by a partial sum of its orthogonal expansion with respect to the basis $\{V_{pq}(x,y)\}$. If the orthogonal moments up to order N are used then the reconstruction formula takes the following form

$$f_N(x,y) = \sum_{p=0}^{N} \sum_{q=0}^{p} \lambda_{p-q,q} V_{p-q,q}(x,y), \tag{1.26}$$

where it was assumed that the orthogonal basis $\{V_{pq}(x,y)\}$ is normalized, i.e., $\iint_\Omega |V_{pq}(x,y)|^2 w(x,y)dxdy = 1$ for all (p,q).

The reconstruction method $f_N(x,y)$ requires the total $(N+1)(N+2)/2$ orthogonal moments and can converge to the true image $f(x,y)$ as N increases. In fact the general theory of orthogonal expansions [206, 217] says that if the image function is square integrable on Ω, i.e, that $\iint_\Omega |f(x,y)|^2 w(x,y)dxdy < \infty$ then due to Parseval's formula we have the following expression for the global reconstruction error

$$\iint_\Omega |f_N(x,y) - f(x,y)|^2 w(x,y)dxdy = \sum_{p=N+1}^{\infty} \sum_{q=0}^{p} |\lambda_{p-q,q}|^2. \tag{1.27}$$

Hence we can readily conclude that the reconstruction error tends to zero, i.e.,

$$\iint_\Omega |f_N(x,y) - f(x,y)|^2 w(x,y)dxdy \to 0 \quad \text{as} \quad N \to \infty. \tag{1.28}$$

Thus we can easily establish the convergence of the reconstruction method $f_N(x,y)$ to any image $f(x,y)$ of the $L_2(\Omega)$ class. The rate at which the error tends to zero depends on the assumed image model. There are several deterministic image models that prevail in the literature. These mathematical models are approximates of real images to a certain degree of faithfulness. The space of functions of bounded variations has had a fundamental impact on image analysis since the

publication of the well-known work of Rudin, Osher, and Fatemi [205]. This image model permits the existence of jumps, edges, and local oscillations in a function and yet it is mathematically tractable. Thus the class of bounded variation functions indeed models faithfully natural images. This is the model which is used throughout this monograph.

An alternate approach for assessing the accuracy of the reconstruction method can be based on establishing the pointwise convergence, i.e., that $f_N(x,y) \to f(x,y)$ as $N \to \infty$ at every point $(x,y) \in \Omega$. This is typically a much more difficult question and can only be resolved for some particular orthogonal basis $\{V_{pq}(x,y)\}$.

Let us finally state that the issue of image reconstruction from object descriptors is crucial since it allows us to visually verify that a sufficient number of features is used to capture the essential structure of the image. This fundamental problem is thoroughly examined in this monograph.

1.5. Discretization and Noise Sensitivity

Thus far we have assumed that one has the ability to observe the image $f(x,y)$ assumed to be a measurable function defined on Ω. In practical applications of image analysis one has to implement any algorithm in the digital domain. Hence the image function $f(x,y)$ is only observed at discrete points. Let $\{(x_i, y_j) : 1 \le i,j \le n\}$ be an $n \times n$ array of pixels, i.e., (x_i, y_j) is the centre point of the (i,j)th pixel. We assume that all pixels are squares with the width Δ, i.e., $x_i - x_{i-1} = y_j - y_{j-1} = \Delta$, and the (i,j)th pixel is defined as the square

$$c_{ij} = [x_i - \Delta/2, x_i + \Delta/2] \times [y_j - \Delta/2, y_j + \Delta/2]. \tag{1.29}$$

As a result of the sampling process applied to the image plane we have only the digitized version $\{f(x_i, y_j), 1 \le i,j \le n\}$ of the image function $f(x,y)$. Consequently one has to reconsider the aforementioned theory of moment invariants for digital images. It is clear that the invariant properties can hold only approximately for $\{f(x_i, y_j), 1 \le i,j \le n\}$ and it is an important issue to obtain an accurate digital approximate of the true moments. This classical problem of numerical integration [31, 40] has been rarely addressed in the context of object invariants, see [117, 144, 220] for some results and Sections 2 and 5 for detailed discussion of this subject.

The simplest approximation (still commonly used in many applications) to compute the geometric moments is the following

$$\widetilde{m}_{pq} = \Delta^2 \sum_{(x_i,y_j)\in\Omega} f(x_i, y_j) x_i^p y_j^q. \tag{1.30}$$

For the orthogonal moments the analogous formula is as follows

$$\widetilde{\lambda}_{pq} = \Delta^2 \sum_{(x_i,y_j)\in\Omega} f(x_i, y_j) V_{pq}(x_i, y_j) w(x_i, y_j). \tag{1.31}$$

Clearly these are not very accurate approximations of m_{pq}, λ_{pq}, respectively. It is worth mentioning that the main focus in the moment invariant literature has been on the development of methods and specialized hardware for fast calculation of the approximations $\widetilde{m}_{pq}$, $\widetilde{\lambda}_{pq}$ [7,11,27,29,45,58,103,114,141,142,156,159,198,208, 236, 246]. The issue of the high precision approximations of the true invariants has been greatly ignored. The first attempt concerning this problem was made in [220] and [144]. In the latter contribution the approximation to m_{pq} was given in the form

$$\overline{m}_{pq} = \sum_{(x_i,y_j)\in\Omega} f(x_i, y_j) \iint_{c_{ij}} x^p y^q dxdy, \tag{1.32}$$

where the factor $\displaystyle\iint_{c_{ij}} x^p y^q dxdy$ can be efficiently evaluated. The errors resulting from these approximations depend not only on the type of the approximation scheme but also on the geometry of the support set Ω. In Sections 2 and 4 we establish a detailed accuracy analysis of various approximation algorithms.

There is yet another important source of error commonly appearing in practical situations, i.e., our original digital image data $\{f(x_i, y_j) : 1 \le i, j \le n\}$ are often observed in the presence of noise. Hence we observe the following noisy version of $f(x_i, y_j)$

$$g(x_i, y_j) = f(x_i, y_j) + z(x_i, y_j), \quad 1 \le i, j \le n, \tag{1.33}$$

where $z(x_i, y_j)$ is a zero mean additive noise process. The noise process $z(x_i, y_j)$ has typically a finite variance and is covariance stationary, i.e., the covariance function of $z(x_i, y_j)$ has the following form

$$\mathrm{cov}(z(x, y), z(x', y')) = \rho\left(\sqrt{(x - x')^2 + (y - y')^2}\right), \tag{1.34}$$

where $\rho(t)$ is an even function.

The issue which arises now is how one can estimate moment descriptors from noisy data. A natural method is to use plug-in approach which replaces the pure digital data $\{f(x_i, y_j) : 1 \le i, j \le n\}$ by the noisy one $\{g(x_i, y_j), 1 \le i, j \le n\}$. This would lead to the following counterparts of (1.30) and (1.31)

$$\widehat{m}_{pq} = \Delta^2 \sum_{(x_i,y_j)\in\Omega} g(x_i, y_j) x_i^p y_j^q \tag{1.35}$$

and

$$\widehat{\lambda}_{pq} = \Delta^2 \sum_{(x_i,y_j)\in\Omega} g(x_i,y_j)V_{pq}(x_i,y_j)w(x_i,y_j), \tag{1.36}$$

respectively.

It is worth noting that both estimates are unbiased versions of $\widetilde{m}_{pq}$ and $\widetilde{\lambda}_{pq}$, i.e., $E\widehat{m}_{pq} = \widetilde{m}_{pq}$ and $E\widehat{\lambda}_{pq} = \widetilde{\lambda}_{pq}$. Having formed the estimates of m_{pq} and λ_{pq} we can consider the problem of image reconstruction from moments. As we have already pointed out the reconstruction of an image from noisy version of the geometric moments is ill-posed and virtually impossible. This is in striking contrast with the orthogonal moments, where the reconstruction algorithm from digital and noisy images is straightforward. Indeed, using (1.26) we can easily form the following reconstruction method

$$\widehat{f}_N(x,y) = \sum_{p=0}^{N}\sum_{q=0}^{p} \widehat{\lambda}_{p-q,q}V_{p-q,q}(x,y), \tag{1.37}$$

for some estimate $\widehat{\lambda}_{pq}$ of λ_{pq}, e.g., like the one given in (1.36).

A simple conclusion which can be drawn from the convergence result in (1.28) is that the more moments are taken into account the more accurate approximation of the true image can be obtained. This is not the case for reconstruction algorithms utilizing the approximated values of λ_{pq}. In fact, by virtue of Parseval's formula applied to $\widehat{f}_N(x,y)$ we obtain the following counterpart of (1.27)

$$\begin{aligned}\iint_\Omega |\widehat{f}_N(x,y) - f(x,y)|^2 w(x,y)dxdy& \\ = \sum_{p=0}^{N}\sum_{q=0}^{p} |\widehat{\lambda}_{p-q,q} - \lambda_{p-q,q}|^2 &+ \sum_{p=N+1}^{\infty}\sum_{q=0}^{p} |\lambda_{p-q,q}|^2. \end{aligned}\tag{1.38}$$

Hence the reconstruction error consists of two parts. The first one is due to the imprecision of estimating the true moments $\{\lambda_{pq}\}$, whereas the second one is a result of using a finite number of moments. Furthermore the first term increases with N, while the second one decreases. This yields a well-known trade-off phenomenon between the estimation and approximation errors. As a result there is an optimal value of N yielding the best optimal reconstruction. This number defines an intrinsic feature vector dimensionality representing an object. It is worth mentioning that the problem of finding the optimal dimensionality of patterns is fundamental in designing classification systems [38].

In Chapters 3 and 4 we establish the asymptotic theory of finding the optimal N for various types of orthogonal moment descriptors. Our theory models the

performance of reconstruction procedures on grids which become increasingly fine, i.e., when $\Delta \to 0$.

The reconstruction algorithm given in (1.37) falls into the category of non-parametric estimation methods since we do not assume any parametric knowledge of a class of image functions taken into account, see [39, 41, 95, 210, 230] for an overview of non-parametric estimation techniques. In our case the parameter N plays the role of the smoothing sequence controlling the aforementioned trade-offs and, as we have already mentioned, has to be selected suitably in order to define a reconstruction method being able to converge to an image function which belongs to some non-parametric function space.

The field of non-parametric estimation has found a great number of applications in science and engineering and this book can be viewed as an extension of this subject to the problem of image analysis. See [80–99, 128, 150–157, 200, 201, 203–254] for the various applications of non-parametric methods in signal processing, system identification, pattern recognition and image processing.

1.6. Conclusions

In this chapter, an overview of moment type image descriptors has been conducted. In particular, we have described the basic concepts of geometric moments, moment invariants, and orthogonal moments. The fundamental issue of recovering an image shape from moments is discussed. This is followed by the examination of the problem of the moment descriptor robustness to discretization and noise.

Moment descriptors are region-based image features as they use all information about the image, i.e., both the image contour and its content unlike the so-called contour-based descriptors such as Fourier descriptors [130, 139] which merely use information about the image boundary. As such they have a number of limitations. First they are sensitive to noise and occlusions because they use a small part of image information. Furthermore in many applications, e.g., face recognition, shape content is more important than the contour. Extensive comparative numerical studies performed in [203, 226] for the problem of off-line recognition of characters and in [251] for the problem of image retrieval suggest that object descriptors based on moments and especially orthogonal moments are the best choices. The methodology developed in this book gives a theoretical support to these empirical findings.

Chapter 2

Image Analysis by Orthogonal Moments

In this chapter a class of moment descriptors stemming from the theory of classical orthogonal polynomials is introduced. We examine basic properties of both orthogonal and geometric moments including the issues of numerical efficiency, reconstruction power from the computed moments, robustness to noise, and automatic selection of optimal number of moments. In particular, the influence of discretization and noise on the accuracy of moment descriptors is thoroughly investigated. Several new numerical techniques that enhance the accuracy and efficiency of orthogonal moment descriptors are proposed. An improved method for calculating the classical geometric moments is also given. We utilize these results for the problem of image reconstruction from orthogonal moments based on a class of classical orthogonal polynomials. The automatic selection of an optimal number of moments is also discussed.

An extension of the above results to the case of generalized moments and orthogonal moments calculated in the digital domain is given.

The main contributions of this chapter are the following:

- An improved estimation technique for high accuracy calculation of geometric moments is proposed.
- Bounds for the precision of the estimates are established.
- An improved technique for high accuracy calculation of orthogonal Legendre moments is proposed.
- The bounds for the precision of the estimates of Legendre moments are derived.

- The issue of a fast calculation of Legendre moments is examined.
- An algorithm for image recovery from Legendre moments is introduced and its asymptotic analysis is investigated.
- Numerical studies are carried out to confirm the basic theory.
- An estimate of geometric moments from noisy images is proposed. The statistical accuracy of the estimates is evaluated.
- An estimate of Legendre moments from noisy images is proposed.
- The reconstruction power of the Legendre moment based image estimate is evaluated. The optimal rate of convergence for the mean integrated squared reconstruction error is established.
- Algorithms for automatic selection of optimal number of moments are proposed.
- The aforementioned theory is further extended to a general class of orthogonal moments.
- Orthogonal moments calculated in the digital domain are discussed.

2.1. Introduction

As was pointed out an essential issue in the field of pattern analysis is the recognition of objects and characters regardless of their positions, sizes, and orientations. Moments and functions of moments have been extensively employed as the invariant global features of an image in pattern recognition, image classification, target identification, and scene analysis, see [160, 199] and the references cited therein.

Generally, these features are invariant under image translation, scale change, and rotation only when they are computed from the original non-distorted analog two-dimensional image. In practice, one observes the digitized, quantized, and often noisy version of the image and the invariance properties are satisfied only approximately. A digital image is obtained by two operations, i.e., by sampling an analog image on a discrete grid followed by the quantization of its amplitude using a finite number of bits. It is worth mentioning that the quantization process can be represented by the image plus noise model [112] and this case is fully covered by the theory developed in this book.

The error analysis and analytic characterization of moment descriptors have been rarely investigated. Some studies concerning the discretization error in the case of geometric moments were performed by Teh and Chin [220], see [117] for

a general theory of digitization errors in computer vision. The vulnerability of moments against noise process has been examined in [1, 173, 221]. In all these studies, only continuous image models have been taken into account, i.e., the discretization error has been greatly ignored.

In this chapter, the detailed analysis of the discretization error occuring in moment computing is carried out. Several new techniques to increase the accuracy and efficiency of moments are proposed. Both the classical geometric moments and orthogonal moments are taken into consideration. Based on the progress made in moment computing, the inverse moment problem of reconstruction of an image from a finite set of moments based on a class of the classical orthogonal polynomials is examined. It is carried out both for discrete and noisy data. This yields a delicate problem of selecting an optimal number of moments from the digital and noisy image. This issue is tackled by cross-validation based selection techniques. Finally we also give a brief introduction to more recently studied discrete orthogonal moments.

2.2. Geometric Moments

Geometric moments are the most commonly used region-based object descriptors, see [160, 199, 203] for an overview of the subject. Let us recall that the geometric moment of order (p, q) is defined as:

$$m_{pq} = \int_{-1}^{+1} \int_{-1}^{+1} x^p \, y^q \, f(x, y) \, dxdy, \tag{2.1}$$

where we assume that the image function $f(x, y)$ is defined on the square $\Omega = [-1, 1]^2$.

If an analog original image function $f(x, y)$ is digitized into its discrete version $\{f(x_i, y_j)\}$ with an $n \times n$ array of pixels, the double integration of (2.1) must be approximated by summation. Here (x_i, y_j) is the centre point of the (i, j) pixel. A commonly used prescription to compute m_{pq} from a digital image is defined as

$$\widetilde{m}_{pq} = \Delta^2 \sum_{i=1}^{n} \sum_{j=1}^{n} x_i^p \, y_j^q \, f(x_i, y_j), \tag{2.2}$$

where $\Delta = x_i - x_{i-1} = y_j - y_{j-1}$ is the sampling interval. A number of fast algorithms and hardware implementations for determining $\widetilde{m}_{pq}$ have been proposed, see, e.g., [24, 29, 141, 199], and [249]. It is clear, however, that $\widetilde{m}_{pq}$ is not a very accurate estimate of m_{pq}, particularly when the moment order (p, q) increases.

The piecewise constant approximation of $f(x, y)$ in (2.1) yields the following approximation of m_{pq}

$$\widehat{m}_{pq} = \sum_{i=1}^{n} \sum_{j=1}^{n} h_{pq}(x_i, y_j)\, f(x_i, y_j), \tag{2.3}$$

where

$$h_{pq}(x_i, y_j) = \int_{x_i - \frac{\Delta}{2}}^{x_i + \frac{\Delta}{2}} \int_{y_j - \frac{\Delta}{2}}^{y_j + \frac{\Delta}{2}} x^p\, y^q\, dxdy \tag{2.4}$$

represents the integration of $x^p y^q$ over the (i, j) pixel.

It is worth noting that if $f(x, y)$ is assumed to be a piecewise constant function over the given pixel set, then

$$\widehat{m}_{pq} = m_{pq}.$$

In practice, however, $f(x, y)$ can be any positive bivariate function and then, clearly, $\widehat{m}_{pq} \neq m_{pq}$.

In the following theorem we evaluate the approximation error $|\widehat{m}_{pq} - m_{pq}|$ for images which are bounded variation (BV) functions on Ω.

The total variation of $f(x, y)$ is defined as

$$\mathrm{TV}(f) = \iint_{\Omega} \left\{ \left| \frac{\partial f(x, y)}{\partial x} \right| + \left| \frac{\partial f(x, y)}{\partial y} \right| \right\} dxdy, \tag{2.5}$$

where the derivatives are meant in the generalized sense [254] . A function $f(x, y)$ with $\mathrm{TV}(f) < \infty$ is said to have bounded variation [60]. The class BV allows the existence of jumps and edges being important features of most natural images. For instance the binary image $\mathbf{1}_G(x, y) \in \mathrm{BV}$, provided that the boundary of the region G is smooth.

Theorem 2.1. *Let the image function $f \in \mathrm{BV}(\Omega)$. Then*

$$|\widehat{m}_{pq} - m_{pq}| \leq \left(\frac{4 f_{max}\, \mathrm{TV}(f)}{(2p+1)(2q+1)} \right)^{\frac{1}{2}} \Delta, \tag{2.6}$$

where $f_{max} = \max_{(x,y)\in\Omega} f(x, y)$.

The proof of this inequality is given in the Appendix.

Inequality (2.6) shows that the error between $\widehat{m}_{pq}$ and m_{pq} decreases as the pixel length Δ becomes smaller. Interestingly, $|\widehat{m}_{pq} - m_{pq}|$ also decreases as the moment order (p, q) increases. The latter property does not hold for the approximation in (2.2) (see Theorem 2.2 in this section).

It is also worth noting that in a general case, $f(x_i, y_j)$ in (2.3) is not the value of $f(x, y)$ at the point (x_i, y_j), but it should be considered as a quantized level of $f(x, y)$ at the pixel (i, j), see [117] for a basic discussion of quantization error in computer vision problems.

The next result describes the error caused by $\widetilde{m}_{pq}$ as an estimate of m_{pq}.

Theorem 2.2. *Let the assumptions of Theorem 2.1 hold. Then for $p, q \geq 1$,*

$$|\widetilde{m}_{pq} - m_{pq}| \leq \frac{f_{max}}{2}(p\Delta + q\Delta) + \left(\frac{4f_{max}\,\mathrm{TV}(f)}{(2p+1)(2q+1)}\right)^{\frac{1}{2}}\Delta. \tag{2.7}$$

The proof of Theorem 2.2 is similar to the proof of Theorem 2.1 and is omitted, see also [144, 173].

It should be noted that the second term on the right-hand side of (2.7) is just the bound for $|\widehat{m}_{pq} - m_{pq}|$. The first term on the right-hand side of (2.7) is an additional bias of $\widetilde{m}_{pq}$ caused by the poorer approximation of the integral in (2.1) and it is an increasing function of the moment order (p, q).

Remark 2.1. *Theorem 2.1 and Theorem 2.2 give the approximation error of $\widehat{m}_{pq}$ and $\widetilde{m}_{pq}$, respectively, under rather general assumption about the image $f(x, y)$. Some stronger conditions imposed on the smoothness of $f(x, y)$ would lead to tighter bounds for the approximation error. For example, for $f(x, y)$ being a piecewise constant function, we can easily obtain that*

$$|\widetilde{m}_{pq} - m_{pq}| \leq C_1\Delta^2,$$

where $C_1 = C_1(p, q)$ is an increasing function of (p, q), and

$$|\widehat{m}_{pq} - m_{pq}| \leq C_2\Delta^2,$$

where C_2 is independent of (p, q).

Furthermore, it is straightforward to observe from (2.4) that

$$\begin{aligned} h_{pq}(x_i, y_j) &= \frac{1}{(p+1)(q+1)}\left[\left(x_i + \frac{\Delta}{2}\right)^{p+1} - \left(x_i - \frac{\Delta}{2}\right)^{p+1}\right] \\ &\quad \cdot \left[\left(y_j + \frac{\Delta}{2}\right)^{q+1} - \left(y_j - \frac{\Delta}{2}\right)^{q+1}\right]. \end{aligned} \tag{2.8}$$

It is clear that the approximation in (2.2) corresponds to

$$h_{pq}(x_i, y_j) = x_i^p y_j^q \Delta^2. \tag{2.9}$$

Applying the binomial expansion to (2.8), one can rewrite $h_{pq}(x_i, y_j)$ as follows

$$\begin{aligned} h_{pq}(x_i, y_j) &= \frac{4}{(p+1)(q+1)} \sum_{l=1}^{p+1} \sum_{t=0}^{q+1} \binom{p+1}{l} \binom{q+1}{t} \\ &\quad \cdot x_i^{p+1-l} y_j^{q+1-t} \left(\frac{\Delta}{2}\right)^{l+t}, \end{aligned} \tag{2.10}$$

where $\sum_{l=1}^{p+1} \sum_{t=0}^{q+1}$ stands for the summation with respect to the odd values of l and t. Plugging this into (2.3), changing variables and recalling the definition of $\widetilde{m}_{pq}$, we obtain

$$\begin{aligned} \widehat{m}_{pq} &= \frac{4}{(p+1)(q+1)} \sum_{l=1}^{p+1} \sum_{t=1}^{q+1} \binom{p+1}{p+1-l} \binom{q+1}{q+1-t} \\ &\quad \cdot \frac{\Delta^{l+t-2}}{2^{l+t}} \widetilde{m}_{p+1-l,q+1-t}. \end{aligned} \tag{2.11}$$

It is interesting to note that the above formula is of the form of the convolution operator. In fact, defining for all odd values of (l, t) the following sequences

$$a_{lt} = \frac{\Delta^{l+t-2}}{2^{l+t}}, \tag{2.12}$$

$$b_{lt} = \binom{p+1}{l} \binom{q+1}{t} \widetilde{m}_{lt}, \tag{2.13}$$

we can rewrite (2.11) as

$$\widehat{m}_{pq} = \frac{4}{(p+1)(q+1)} \{a_{lt}\} * \{b_{lt}\}, \tag{2.14}$$

where $*$ is the two-dimensional $(p+1)(q+1)$-point discrete convolution operator. The representation in (2.14) allows us to adapt a number of existing fast computation techniques for $\{\widetilde{m}_{lt}\}$, see [24, 29, 141, 199, 249], to our modified moments $\widehat{m}_{pq}$.

The algebraic moment invariants, due originally to Hu [101], see also [203], require the knowledge of all m_{pq}'s for $p + q \leq 3$. Owing to (2.11), we can obtain

the following list of relationships between $\{\widetilde{m}_{pq}\}$ and $\{\widehat{m}_{pq}\}$ for $p+q \leq 3$

$$\widehat{m}_{00} = \widetilde{m}_{00}; \qquad \widehat{m}_{10} = \widetilde{m}_{10}; \qquad \widehat{m}_{01} = \widetilde{m}_{01};$$
$$\widehat{m}_{20} = \widetilde{m}_{20} + \frac{\Delta^2}{12}\widetilde{m}_{00}; \qquad \widehat{m}_{02} = \widetilde{m}_{02} + \frac{\Delta^2}{12}\widetilde{m}_{00}; \qquad \widehat{m}_{11} = \widetilde{m}_{11};$$
$$\widehat{m}_{30} = \widetilde{m}_{30} + \frac{\Delta^2}{4}\widetilde{m}_{10}; \qquad \widehat{m}_{21} = \widetilde{m}_{21} + \frac{\Delta^2}{12}\widetilde{m}_{01}; \qquad \widehat{m}_{12} = \widetilde{m}_{12} + \frac{\Delta^2}{12}\widetilde{m}_{10};$$
$$\widehat{m}_{03} = \widetilde{m}_{03} + \frac{\Delta^2}{4}\widetilde{m}_{01}.$$

Furtermore, the first three leading terms in (2.11) are

$$\widehat{m}_{pq} = \widetilde{m}_{pq} + \frac{q(q-1)}{24}\Delta^2\widetilde{m}_{p,q-2} + \frac{p(p-1)}{24}\Delta^2\widetilde{m}_{p-2,q} + O(\Delta^4). \tag{2.15}$$

This again explains that the approximation error resulting in the use of $\widetilde{m}_{pq}$'s quickly gets out of control when the moment order increases. Obviously, when the moments of higher order are involved, the problem of the accurate numerical approximation must be addressed prior to any efficient numerical or hardware implementations.

The above discussion forms the basis for our studies concerning the digital approximation for orthogonal moment descriptors. We use orthogonal moments due to the fact that they possess better reconstruction power than geometrical moments and they attain a zero value of redundancy measures [1,173,219,221]. In particular, Legendre and Gegenbauer orthogonal moments are examined. We use both types of moments due to their direct connection to the geometric moments. In the case of Gegenbauer orthogonal moments we can obtain a further refinement in the accuracy by selecting a scaling factor appearing in the definition of the Gegenbauer polynomials. The methodology presented can also be extended to other types of moments known in the literature [219, 221].

2.3. Legendre Moments

The Legendre moment of order (p, q) is defined as

$$\lambda_{pq} = \int_{-1}^{+1}\int_{-1}^{+1} P_p(x)\, P_q(y)\, f(x,y)dxdy, \tag{2.16}$$

where the p-th order Legendre polynomial is given by the Rodrigues formula

$$P_p(x) = \frac{1}{2^p p!}\frac{d^p}{dx^p}(x^2-1)^p, \quad x \in [-1,1].$$

See [206], for basic properties of the Legendre polynomials.

Hence, λ_{pq} generalizes m_{pq} in the sense that the monomial $x^p y^q$ is replaced by the orthogonal polynomial $P_p(x)P_q(y)$ of the same order. In Fig. 2.1 we depict a collection of the Legendre polynomials $\{P_p(x)P_q(y)\}$ for $(p,q) \in \{(2,0),(0,2),(2,1);(4,0),(2,2),(4,4)\}$.

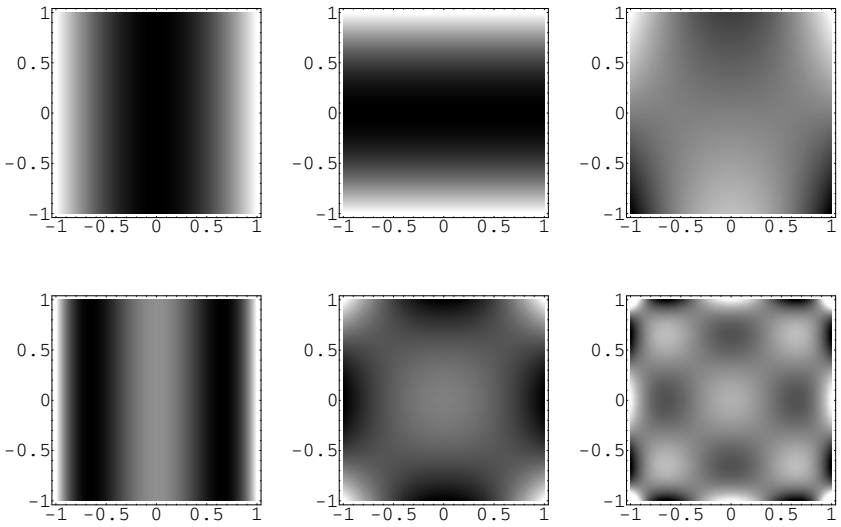


Figure 2.1: Density plots of $P_p(x)P_q(y)$ for $(p,q) \in \{(2,0),(0,2),(2,1);(4,0),(2,2),(4,4)\}$

Similarly to the case of the geometric moments, we can approximate λ_{pq} by

$$\widehat{\lambda}_{pq} = \sum_{i=1}^{n}\sum_{j=1}^{n} h_{pq}(x_i, y_j)\, f(x_i, y_j), \tag{2.17}$$

where

$$h_{pq}(x_i, y_j) = \int_{x_i-\frac{\Delta}{2}}^{x_i+\frac{\Delta}{2}} \int_{y_j-\frac{\Delta}{2}}^{y_j+\frac{\Delta}{2}} P_p(x)\, P_q(y)\, dxdy. \tag{2.18}$$

The equation in (2.18) is a counterpart of (2.4). It is common, however, in the computer vision literature to use a simpler approximation of λ_{pq} of the form

$$\widetilde{\lambda}_{pq} = \Delta^2 \sum_{i=1}^{n}\sum_{j=1}^{n} P_p(x_i)P_q(y_j)f(x_i, y_j). \tag{2.19}$$

The approximation properties of $\widetilde{\lambda}_{pq}$ are similar to $\widetilde{m}_{pq}$ in (2.2), i.e., $\widetilde{\lambda}_{pq}$ is inferior to $\widehat{\lambda}_{pq}$. The following theorem describes a result concerning the accuracy of $\widehat{\lambda}_{pq}$.

Theorem 2.3. *Let the assumptions of Theorem 2.1 hold. Then*

$$|\widehat{\lambda}_{pq} - \lambda_{pq}| \leq \left(\frac{4}{(2p+1)(2q+1)} f_{max}\, \mathrm{TV}(f) \right)^{\frac{1}{2}} \Delta. \tag{2.20}$$

The proof of this theorem is similar to the proof of Theorem 2.1 and therefore it is omitted, see also [144, 173].

It should also be remarked that for a class of piecewise constant image functions over Ω, we can show that

$$|\widehat{\lambda}_{pq} - \lambda_{pq}| \leq \alpha \frac{1}{\sqrt{pq}} \Delta^2, \tag{2.21}$$

for some constant α being independent of (p, q) and Δ. This should be compared with the result in Remark 2.1.

Although we can obtain an explicit formula for $h_{pq}(x_i, y_j)$ in (2.18), analogously as in (2.8) (see the discussion below) we wish to use some techniques of numerical integration [31] since for other orthogonal moments such formulas are not feasible. In particular, the alternative extended Simpson's rule is used in our numerical examples to approximate the integral in (2.18) [31].

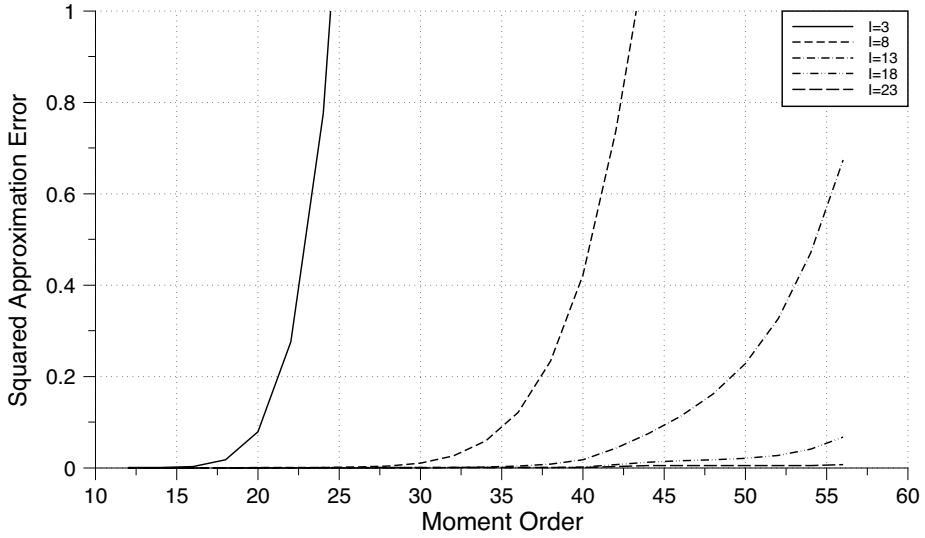


Figure 2.2: The error E versus the moment order for five different numerical integration rules applied to a constant image

To show a potential benefit from the use of such numerical techniques, we assume that the image function $f(x, y)$ for all (x, y) has the same constant value

a. In such a case, all Legendre moments should be equal to zero except $\lambda_{00} = 4a$. We use the sum of all squared Legendre moments up to the order $p + q \leq N_{max}$ (except $p = q = 0$) as the measure of approximation error, i.e.,

$$E = \sum_{p=0}^{N_{max}} \sum_{q=0}^{p} \widehat{\lambda}_{p-q,q}^2 - \widehat{\lambda}_{00}^2. \tag{2.22}$$

Clearly, the smaller the E value in (2.22), the better the performance of the approximation. Five different numerical integration rules, $I = 3, I = 8, I = 13, I = 18$, and $I = 23$ are employed. The factor I stands for the number of points within a given pixel required to evaluate the integral in (2.18). All E's are listed in Table 2.1 and illustrated in Fig. 2.2. The highest Legendre moment order used in this experiment is 56, i.e., $N_{max} = 56$.

Only the E's which are less than 1.0 are presented in Fig. 2.2. Each E increases very sharply after the moment order is over a certain number. As expected, the higher accuracy approximation rules perform better than the lower ones.

2.3.1. Numerical Efficiency

As we have already pointed out, when the higher order Legendre polynomials $\{P_p(x)P_q(y)\}$ are involved we need an efficient method for generating $\{P_p(x)\}$. For small values of p, the three-term recurrence formula [206]

$$P_{p+1}(x) = \frac{2p+1}{p+1} x P_p(x) - \frac{p}{p+1} P_{p-1}(x) \tag{2.23}$$

with $P_0(x) = 1$ and $P_1(x) = x$ can be used. For larger values of p and larger images, however, we have avoided (2.23) and our computations have been based on the look-up table approach, i.e., the Legendre polynomials up to the order 55 have been generated. Alternatively, we can employ the asymptotic approximations for $P_p(x)$, e.g., the first order Laplace's formula [206, 217]

$$P_p(\cos\theta) = \sqrt{2}(\pi p \sin\theta)^{-1/2} \cos((p + 1/2)\theta - \pi/4) + O(p^{-3/2}),$$

for $0 < \theta < \pi$.

Numerical techniques for generating orthogonal polynomials can be found in [57].

The discussion in the previous section reveals that the Legendre moments of the digital image $\{f(x_i, y_j)\}$ can be obtained from the formulas in (2.17) and

Table 2.1: The approximation error E for different integration rules indexed by I being the number of points which are equally spaced apart inside a single pixel

Order	$I = 3$	$I = 8$	$I = 13$	$I = 18$	$I = 23$
12	0.00003				
14	0.00039				
16	0.00308				
18	0.01780				
20	0.07873	0.00002			
22	0.27537	0.00009			
24	0.77483	0.00036	0.00001		
26	1.77074	0.00122	0.00002		
28		0.00373	0.00008	0.00001	
30		0.01027	0.00024	0.00002	
32		0.02562	0.00064	0.00005	0.00001
34		0.05822	0.00160	0.00012	0.00002
36		0.12126	0.00372	0.00029	0.00004
38		0.23284	0.00807	0.00068	0.00010
40		0.42134	0.01779	0.00190	0.00043
42		0.73673	0.04279	0.00728	0.00294
44		1.15393	0.07487	0.01208	0.00436
46			0.11278	0.01524	0.00458
48			0.16089	0.01773	0.00459
50			0.22783	0.02100	0.00469
52			0.32621	0.02745	0.00469
54			0.47053	0.04105	0.00499
56			0.67317	0.06711	0.00695

(2.18). This can be done by numerical integration or they can be even calculated explicitly. In fact, employing the identity [206]

$$\int_{-1}^{x} P_p(u)du = \frac{1}{2p+1}\{P_{p+1}(x) - P_{p-1}(x)\}$$

we can easily evaluate the factor $h_{pq}(x_i, y_j)$ in (2.18) as follows

$$\begin{aligned} h_{pq}(x_i, y_j) &= \frac{1}{(2p+1)(2q+1)}\{(P_{p+1}(x_i + \Delta/2) \\ &\quad - P_{p+1}(x_i - \Delta/2))(P_{p-1}(x_i + \Delta/2) \\ &\quad - P_{p-1}(x_i - \Delta/2))\}\{(P_{q+1}(y_j + \Delta/2) \\ &\quad - P_{q+1}(y_j - \Delta/2))(P_{q-1}(y_j + \Delta/2) - P_{q-1}(y_j - \Delta/2))\}. \end{aligned}$$

2.4. Image Reconstruction from Moments

In the previous sections, several problems of accuracy and efficiency in computing individual moments have been addressed. To verify their global properties, the problem of the image reconstruction from moments should be examined. As has been pointed out, however, the recovery of an image from the geometric moments is strongly ill-posed and computationally expensive [218]. On the other hand, the reconstruction from orthogonal moments does not need regularization, it adds the individual components of each order to generate the reconstructed image. Hence, in this section, we discuss the problem of the image reconstruction from the Legendre moment descriptors.

A problem which is raised here can be stated as follows: if only a finite set of moments of an image are given, how well can we reconstruct the image?

The Legendre polynomials $\{P_p(x)\}$ constitute a complete orthogonal set on the interval $[-1, 1]$, i.e.,

$$\int_{-1}^{+1} P_p(x)\, P_q(x)\, dx = \frac{2}{2p+1}\, \delta_{pq}, \tag{2.24}$$

where δ_{pq} is the Kronecker function, i.e., $\delta_{pq} = 1$ if $p = q$ and 0 otherwise.

Consequently it can be shown [126] that $\{P_p(x)P_q(y)\}$ is an orthogonal basis on $\Omega = [-1, 1]^2$. Hence provided that $f(x, y)$ is in the $L_2(\Omega)$ space, we can represent the image function $f(x, y)$ as a formal infinite series expansion

$$f(x, y) = \sum_{p=0}^{\infty} \sum_{q=0}^{p} \tau_{p-q,q}\, \lambda_{p-q,q}\, P_{p-q}(x)\, P_q(y), \tag{2.25}$$

where the Legendre moment λ_{pq} of $f(x, y)$ of order (p, q) is defined in (2.16) and

$$\tau_{pq} = \frac{(2p+1)(2q+1)}{4} \tag{2.26}$$

is the normalizing sequence due to (2.24).

In practice, however, one has to truncate infinite series in (2.25). If only the Legendre moments of order smaller than N_{max} are given, the function $f(x, y)$ can be approximated by a truncated series

$$f(x, y) \simeq f_{N_{max}}(x, y) = \sum_{p=0}^{N_{max}} \sum_{q=0}^{p} \tau_{p-q,q}\, \lambda_{p-q,q}\, P_{p-q}(x)\, P_q(y). \tag{2.27}$$

Furthermore, λ_{pq}'s must be replaced by their approximations given by (2.17), yielding the following reconstruction scheme

$$\widehat{f}_{N_{max}}(x,y) = \sum_{p=0}^{N_{max}} \sum_{q=0}^{p} \tau_{p-q,q}\, \widehat{\lambda}_{p-q,q}\, P_{p-q}(x)\, P_q(y). \tag{2.28}$$

This is actually the basic equation used in the image reconstruction via the Legendre moments. It is important to note that when the given order N_{max} is to be increased, the previously determined $\widehat{\lambda}_{pq}$'s do not change.

2.4.1. Reconstruction Error Analysis

To measure the error between the original image and its reconstructed version, the following global L_2 distance is employed

$$\text{Error}(\widehat{f}_{N_{max}}) = \int_{-1}^{1}\int_{-1}^{1} [\widehat{f}_{N_{max}}(x,y) - f(x,y)]^2 dxdy, \tag{2.29}$$

where N_{max} is the number of moments involved in reconstruction, and $\widehat{f}_{N_{max}}(x,y)$ represents the reconstructed image from $f(x,y)$. Let us note that $\text{Error}(\widehat{f}_{N_{max}})$ represents the energy of the error image $\widehat{f}_{N_{max}}(x,y) - f(x,y)$.

Due to (2.28) and (2.25), we obtain

$$\begin{aligned} \widehat{f}_{N_{max}}(x,y) - f(x,y) &= \sum_{p=0}^{N_{max}} \sum_{q=0}^{p} \tau_{p-q,q}\, [\widehat{\lambda}_{p-q,q} - \lambda_{p-q,q}]\, P_{p-q,q}(x)\, P_q(y) \\ &\quad - \sum_{p=N_{max}+1}^{\infty} \sum_{q=0}^{p} \tau_{p-q,q} \lambda_{p-q,q} P_{p-q,q}(x) P_q(y). \end{aligned} \tag{2.30}$$

Then, by virtue of Parseval's formula and (2.24), we have

$$\begin{aligned} \text{Error}(\widehat{f}_{N_{max}}) &= \sum_{p=0}^{N_{max}} \sum_{q=0}^{p} \tau_{p-q,q} [\widehat{\lambda}_{p-q,q} - \lambda_{p-q,q}]^2 \\ &\quad + \sum_{p=N_{max}+1}^{\infty} \sum_{q=0}^{p} \tau_{p-q,q} \lambda_{p-q,q}^2. \end{aligned} \tag{2.31}$$

As is seen from (2.31), the reconstruction error $\text{Error}(\widehat{f}_{N_{max}})$ consists of two parts. The first part is due to the discrete approximation of the true moments $\{\lambda_{pq}\}$, while the second one is a result of using a finite number of moments. It is clear

that the first term in (2.31) increases as $N_{max} \to \infty$, whereas the second one tends to zero as $N_{max} \to \infty$.

With the improved moment methods introduced and examined in the previous sections, we can reduce the discrete approximation error (quantified by the first term in (2.31)) to a tolerable low level. The experimental results of image reconstruction via Legendre moments, which will be presented in the next section, will indicate that when the maximum given order N_{max} reaches a certain value, $\widehat{f}_{N_{max}}(x, y)$ can be very close to the original image function $f(x, y)$.

2.4.2. Experimental Results

In the experiments, a set of five Chinese characters, shown in Fig. 2.3, are used as the test images. Each image consists of 24×24 pixels and the range of graylevels for each pixel is 32. All characters have the gray level 11 and the background has the value 21. The normalized mean square error between the original image

Figure 2.3: Five original Chinese characters used in image reconstruction via the Legendre moments. From left to right are C_1, C_2, C_3, C_4, and C_5

$f(x, y)$ and the reconstructed image $\widehat{f}_{N_{max}}(x, y)$ is defined by

$$\begin{aligned} e^2_{N_{max}} &= \frac{\text{Error}(\widehat{f}_{N_{max}})}{\iint_\Omega [f(x,y)]^2 dxdy} \\ &= \frac{\iint_\Omega [\widehat{f}_{N_{max}}(x,y) - f(x,y)]^2 dxdy}{\iint_\Omega [f(x,y)]^2 dxdy}, \end{aligned} \tag{2.32}$$

which is considered as a measure of the image reconstruction ability.

The alternative extended Simpson's rule with order $I = 23$ is applied to compute the Legendre moments in this experiment. Table 2.2 and Fig. 2.6 show the $e^2_{N_{max}}$ values from the reconstructed Chinese characters from order 2 up to order 56. It should be noted that the $e^2_{N_{max}}$ decreases monotonically in the cases of all five characters.

Table 2.2: The values of normalized reconstruction errors for the five reconstructed Chinese characters.

Order	C_1	C_2	C_3	C_4	C_5
2	0.047615	0.045822	0.046409	0.043936	0.044973
4	0.046324	0.045556	0.045387	0.043429	0.043358
6	0.045428	0.043609	0.044421	0.041650	0.042230
8	0.042815	0.040956	0.043409	0.041143	0.040772
10	0.040362	0.039399	0.041322	0.038537	0.039932
12	0.038164	0.037170	0.037279	0.036748	0.037710
14	0.032964	0.034942	0.034077	0.033165	0.034441
16	0.030828	0.032760	0.030799	0.030356	0.032472
18	0.027089	0.029455	0.028143	0.026889	0.029127
20	0.024840	0.025035	0.025198	0.023493	0.026409
22	0.021336	0.021264	0.022372	0.021577	0.023423
24	0.017141	0.019605	0.019870	0.019955	0.020928
26	0.014513	0.016060	0.017691	0.018034	0.018245
28	0.012146	0.012868	0.014819	0.015282	0.015386
30	0.010568	0.010343	0.012224	0.012753	0.012617
32	0.008775	0.008540	0.009367	0.010563	0.010127
34	0.007346	0.007377	0.007456	0.008500	0.008573
36	0.006547	0.006526	0.006485	0.007162	0.007298
38	0.005348	0.005645	0.005668	0.006265	0.006336
40	0.004564	0.004769	0.004799	0.005378	0.004980
42	0.003996	0.004293	0.004219	0.004653	0.004322
44	0.003504	0.003734	0.003576	0.004111	0.003786
46	0.003217	0.003181	0.003165	0.003584	0.003199
48	0.003048	0.002869	0.002811	0.003041	0.002865
50	0.002607	0.002613	0.002728	0.002691	0.002647
52	0.002408	0.002321	0.002394	0.002605	0.002384
54	0.002408	0.002556	0.002451	0.002332	0.002419
56	0.002377	0.002602	0.002415	0.002271	0.002176

Fig. 2.4 shows the original Chinese character C_1 and its reconstructed patterns. The first image on the first row illustrates the original image of the character C_1. The second column to the ninth column of the first row display its reconstructed patterns with order 6, 8, 10, 12, 14, 16, 18 and 20; the second and third rows show the reconstructed character C_1 with order 22 up to 56, respectively.

Fig. 2.5 shows the five original Chinese characters and their reconstructed patterns from the higher order Legendre moments. The first column illustrates

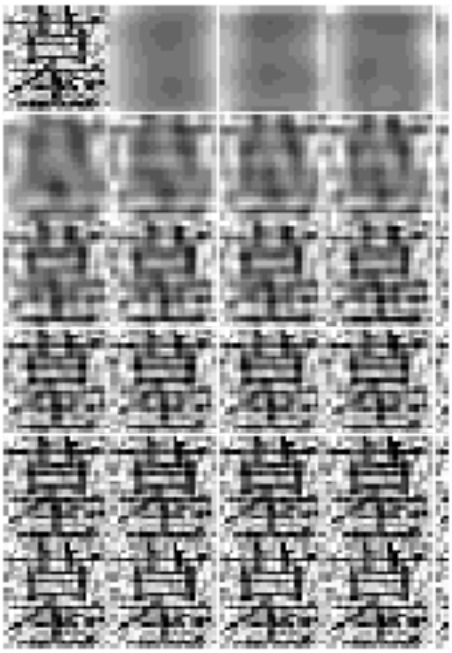


Figure 2.4: The Chinese character C_1 and its reconstructed patterns based on the Legendre moments

five original characters. The second column to the ninth column display the reconstructed patterns of all characters in the first column with order 28, 32, 36, 40, 44, 48, 52 and 56, respectively. Clearly, the numerical results shown in Table 2.2 and Fig. 2.6 are concordant with the visual results presented in Fig. 2.5.

2.5. Moment Descriptors for Noisy Images

In the previous sections we have examined the accuracy of moment descriptors computed from discrete, noise-free data. In most practical situations, however, one observes noisy version of an image and would like to characterize the original image via moments.

Hence, our aim in this section is to study properties of moments from noisy data

$$g(x_i, y_j) = f(x_i, y_j) + z(x_i, y_j), \tag{2.33}$$

$1 \leq i, j \leq n$, where $z(x, y)$ is zero mean, uncorrelated random error with a finite variance σ^2. We use the uncorrelated noise model in order to simplify

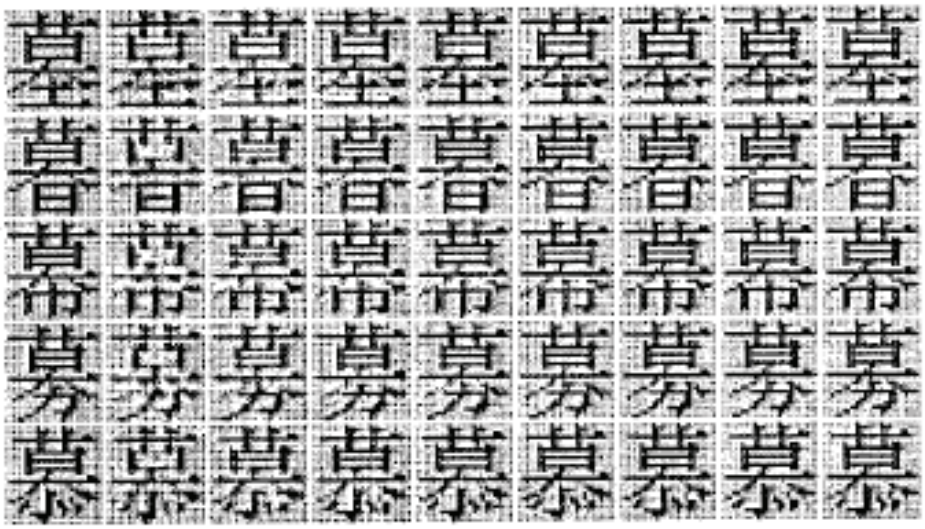


Figure 2.5: Five Chinese characters and their reconstructed patterns utilizing the Legendre moments.

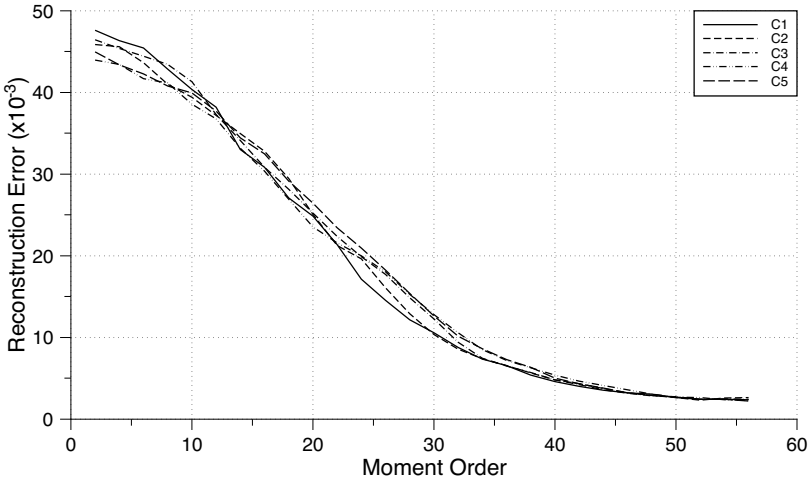


Figure 2.6: Normalized reconstruction errors for the five reconstructed Chinese characters

our mathematical derivations. The extension to a noise model of the spatially stationary structure is possible and leads to similar conclusions.

The geometric moments $\widetilde{m}_{pq}$ and $\widehat{m}_{pq}$ are now defined as in (2.2) and (2.3) with $f(x_i, y_j)$ replaced by $g(x_i, y_j)$.

Hence we have

$$\widetilde{m}_{pq} = \Delta^2 \sum_{i=1}^{n} \sum_{j=1}^{n} x_i^p \, y_j^q \, g(x_i, y_j), \tag{2.34}$$

$$\widehat{m}_{pq} = \sum_{i=1}^{n} \sum_{j=1}^{n} h_{pq}(x_i, y_j) \, g(x_i, y_j), \tag{2.35}$$

where $h_{pq}(x_i, y_j)$ is given in (2.4).

To verify the robustness of $\widetilde{m}_{pq}$, $\widehat{m}_{pq}$ against the random additive noise, let us evaluate the variance of both approximations. This is presented in the following theorem.

Theorem 2.4. *Let the image function $f(x, y)$ be sampled according to (2.33). Then*

$$\mathrm{var}(\widehat{m}_{pq}) \leq \frac{4\sigma^2}{(2p+1)(2q+1)} \Delta^2, \tag{2.36}$$

$$\mathrm{var}(\widetilde{m}_{pq}) \leq \sigma^2 \left(\frac{4}{(2p+1)(2q+1)} + p\Delta + q\Delta \right) \Delta^2. \tag{2.37}$$

The proof of Theorem 2.4 can be found in [144, 173].

From the above result it follows that $\widetilde{m}_{pq}$ is more vulnerable to noise than $\widehat{m}_{pq}$.

Remark 2.2. *Using the technique similar as in the proof of Theorem 2.4, one can also show that both* $\mathrm{cov}(\widetilde{m}_{pq}, \widetilde{m}_{rs})$ *and* $\mathrm{cov}(\widehat{m}_{pq}, \widehat{m}_{rs})$ *are of order*

$$\sigma^2 \frac{(1-(-1)^{p+r+1})(1-(-1)^{q+s+1})}{(p+r+1)(q+s+1)} \Delta^2 \tag{2.38}$$

as $\Delta \to 0$.

The result in (2.38) is, however, asymptotic in nature, whereas Theorem 2.4 gives the bounds for any Δ.

As for the Legendre moments, the formula in (2.17) takes now the following form

$$\widehat{\lambda}_{pq} = \sum_{i=1}^{n} \sum_{j=1}^{n} h_{pq}(x_i, y_j) g(x_i, y_j), \tag{2.39}$$

with $h_{pq}(x_i, y_j)$ defined in (2.18).

Using the results of [144,173], we can obtain the following bounds concerning the variance of $\widehat{\lambda}_{pq}$.

Theorem 2.5. *Let the conditions of Theorem 2.4 hold. Then*

$$\mathrm{var}(\widehat{\lambda}_{pq}) \leq \frac{4\sigma^2}{(2p+1)(2q+1)}\Delta^2. \tag{2.40}$$

Remark 2.3. *It can also be demonstrated that*

$$\mathrm{cov}(\widehat{\lambda}_{pq}, \widehat{\lambda}_{rs}) \simeq \frac{4\sigma^2}{(2p+1)(2q+1)}\Delta^2\delta_{pr}\delta_{qs} \tag{2.41}$$

as $\Delta \to 0$.

This result shows a striking difference between the geometric and orthogonal moments as far as the noise robustness is concerned. The orthogonal moments are uncorrelated, whereas the geometric moments are correlated.

It is also worth noting that the results of Theorem 2.5 and Remark 2.3 should be contrasted with those obtained in [221] where the variance of $\widehat{\lambda}_{pq}$ is an increasing function of the moment order. Such a difference is due to a normalization used in [173] for defining λ_{pq} which gives preference to other moment descriptors.

Taking Remark 2.2 and Remark 2.3 into account, we can conclude that the variability of $\widehat{m}_{pq}$ and $\widehat{\lambda}_{pq}$ seems to be identical. Nevertheless, as has already been pointed out the reconstruction techniques based on the orthogonal moments outperform those using the geometric moment descriptors. The reconstruction issue in the case of noisy and discrete data is examined in the next section.

2.6. Image Reconstruction from Noisy Data

Taking (2.28) and (2.39) into account, it is natural to define reconstruction technique from the Legendre moments as follows

$$\widehat{f}_{N_{max}}(x, y) = \sum_{p=0}^{N_{max}} \sum_{q=0}^{p} \tau_{p-q,q}\widehat{\lambda}_{p-q,q}P_{p-q}(x)P_q(y), \tag{2.42}$$

where $\widehat{\lambda}_{pq}$ is given in (2.39).

As the performance measure for $\widehat{f}_{N_{max}}(x, y)$, we use the integrated squared error $\mathrm{Error}(\widehat{f}_{N_{max}})$ given in (2.29).

Proceeding as in (2.31), we can easily obtain

$$\begin{aligned}\mathrm{Error}(\widehat{f}_{N_{max}}) &= \sum_{p=0}^{N_{max}}\sum_{q=0}^{p}\tau_{p-q,q}(\widehat{\lambda}_{p-q,q}-\lambda_{p-q,q})^2 \\ &\quad + \sum_{p=N_{max+1}}^{\infty}\sum_{q=0}^{p}\tau_{p-q,q}\lambda_{p-q,q}^2. \end{aligned} \tag{2.43}$$

Since $\widehat{\lambda}_{pq}$ is a random variable, let us evaluate the expected value of $\mathrm{Error}(\widehat{f}_{N_{max}})$. It is plain that this requires the evaluation the term $(\widehat{\lambda}_{p-q,q}-\lambda_{p-q,q})^2$ in (2.43).

It is clear that

$$E(\widehat{\lambda}_{p-q,q}-\lambda_{p-q,q})^2 = \mathrm{var}(\widehat{\lambda}_{p-q,q}) + (E\widehat{\lambda}_{p-q,q}-\lambda_{p-q,q})^2.$$

This yields the following decomposition of the expected value of the error

$$\begin{aligned}E(\mathrm{Error}(\widehat{f}_{N_{max}})) &= \sum_{p=0}^{N_{max}}\sum_{q=0}^{p}\tau_{p-q,q}\,\mathrm{var}(\widehat{\lambda}_{p-q,q}) \\ &\quad + \sum_{p=0}^{N_{max}}\sum_{q=0}^{p}\tau_{p-q,q}(E\widehat{\lambda}_{p-q,q}-\lambda_{p-q,q})^2 + \sum_{p=N_{max}+1}^{\infty}\sum_{q=0}^{p}\tau_{p-q,q}\lambda_{p-q,q}^2. \end{aligned} \tag{2.44}$$

Every component in the decomposition is due to different type of distortions present in the image. The first term in (2.44) is merely caused by the noise process. The second term is due to the discretization error and it has already been examined in Section 3. Finally, the last term in (2.44) represents the truncation error caused by the fact that our reconstruction technique in (2.42) uses the moments up to the order N_{max} only.

Recalling the results of Theorem 2.3, Theorem 2.5 and the fact that the total number of moments used in (2.42) is $(N_{max}+1)(N_{max}+2)/2$, we can easily evaluate the first two terms in (2.44). This yields the following result.

Theorem 2.6. *Let the assumptions of Theorem 2.3 and Theorem 2.5 hold. Then*

$$\begin{aligned}E(\mathrm{Error}(\widehat{f}_{N_{max}})) &\le \sigma^2\frac{(N_{max}+1)(N_{max}+2)}{2}\Delta^2 \\ &\quad + f_{max}\,\mathrm{TV}(f)\frac{(N_{max}+1)(N_{max}+2)}{2}\Delta^2 + \mathrm{TR}(N_{max}), \end{aligned} \tag{2.45}$$

where $\mathrm{TR}(N_{max})$ *represents the last term in (2.44).*

Remark 2.4. *The first two terms in (2.45) are of the same order. This is due to the fact that we employ the bound given in Theorem 2.3. For smooth images, see Remark 2.1 and (2.21), the second term in (2.45) will be of order* $O((N_{max}\Delta^2)^2)$.

Let us note again that $\mathrm{TR}(N_{max}) \to 0$ as $N_{max} \to \infty$. This reveals an apparent trade-off in (2.45) since the first two terms tend to infinity as $N_{max} \to \infty$. Hence, these two factors are working against each other indicating that the expected reconstruction error $E(\mathrm{Error}(\widehat{f}_{N_{max}}))$ will initially decrease as N_{max} is increasing to a certain optimal value N^*_{max} and then it will increase to infinity. A specific value of N^*_{max} depends on the speed at which $\mathrm{TR}(N_{max})$ tends to zero. This rate is controlled by the smoothness of the image function $f(x, y)$. To illustrate this let $f(x, y)$ be a pure step edge image along the y-axis with the values f_1 and f_2 on opposite sides of the edge. An involved but straightforward algebra yields

$$\mathrm{TR}(N_{max}) \leq \frac{(f_1 - f_2)^2}{2\pi} \frac{1}{N_{max}}. \tag{2.46}$$

One can generalize this example to the case of bounded variation functions. Indeed, using the fact [206]

$$\left| \int_{-1}^{x} P_p(u)du \right| \leq \frac{4}{\sqrt{\pi}(p+1)^{3/2}}$$

for $|x| \leq 1$ and by some algebra we can obtain the following bound for the Legendre moment

$$|\lambda_{pq}| \leq c(p+1)^{-3/2}(q+1)^{-3/2},$$

where c depends on $\mathrm{TV}(f)$ - the total variation of $f(x, y)$.

All these considerations and some algebra give the following rate of decay of the approximation error.

Theorem 2.7. *Let the image function* $f \in \mathrm{BV}(\Omega)$. *Then*

$$\mathrm{TR}(N_{max}) = O\left(\frac{1}{N_{max}}\right).$$

Combining this result with the result of Theorem 2.6 we can readily obtain the following result on the accuracy of the reconstruction algorithm.

Theorem 2.8. *Let the image function* f *satisfy the conditions of Theorem 2.6 and Theorem 2.7. Then selecting*

$$N^*_{max} = \alpha\Delta^{-\frac{2}{3}},$$

we obtain

$$E\ \text{Error}(\widehat{f}_{N^*_{max}}) = O(\Delta^{\frac{2}{3}}). \tag{2.47}$$

Remark 2.5. *The rate of convergence derived in (2.47) is relatively slow. This is due to both the richness of the class of bounded variation functions and the two-dimensional nature of the image function $f(x, y)$. The former requirement limits the decay of the truncation error, whereas the latter one makes the variance term larger. Faster rates can be obtained either by increasing the smoothness of the image function $f(x, y)$ or using some one-dimensional approximations of $f(x, y)$. The first choice is rather unacceptable in image analysis since images are inherently discontinuous due to the presence of edges. Regarding the second alternative we can, e.g., represent $f(x, y)$ by an additive function, i.e., $f(x, y) = f_1(x) + f_2(y)$, for some single-variable functions $f_1(x), f_2(y)$. Then using the results of this chapter and some methodology borrowed from [97] we can show that the reconstruction error is of order $O(\Delta)$ for all bounded variation images of the additive form. This is clearly much faster rate than that given in (2.47). Yet another interesting avenue could be based on the remarkable theorem of Kolmogorov and Arnold [151] which says that every continuous bivariate function on the square $[-1, 1]^2$ is a sum of at most five functions of the following form $f(x, y) = g(u(x)+v(y))$, where $g(\cdot), u(\cdot), v(\cdot)$ are univariate continuous functions.*

This representation for bivariate continuous functions combined with the decomposition of the bounded variation functions into the continuous part and the pure edge part would yield "building block" for powerful univariate representations of images. This issue is left for future research.

Since the constant α appearing in Theorem 2.8 depends on the unknown image, the proposed N^*_{max} cannot be used in practice. A practical problem arises how to evaluate N^*_{max} directly from the available data. This issue is discussed in the next section.

In order to verify the aforementioned properties, the Chinese character C_1 is employed as the testing pattern in our experiments. Fig. 2.7 shows the trend of the integrated squared reconstruction error $\text{Error}(\widehat{f}_{N_{max}})$ averaged over 10 runs. The noise process is assumed to be Gaussian with $\sigma^2 = 4.0$. As expected, the error first decreases, reaches minimum at $N_{max} = 35$ and then increases. Table 2.3 lists the numerical values of $\text{Error}(\widehat{f}_{N_{max}})$ which is named as $I(N)$. Fig. 2.8 illustrates the noisy version of the Chinese character C_1 and its reconstructed images from order 4 up to order 56, from left to right, first row to last row, respectively.

From equation (2.45), it can be easily concluded that when the level of noise increases, the optimal number of moments N^*_{max} minimizing the reconstruction error becomes smaller. An experiment was designed to obtain further insight into

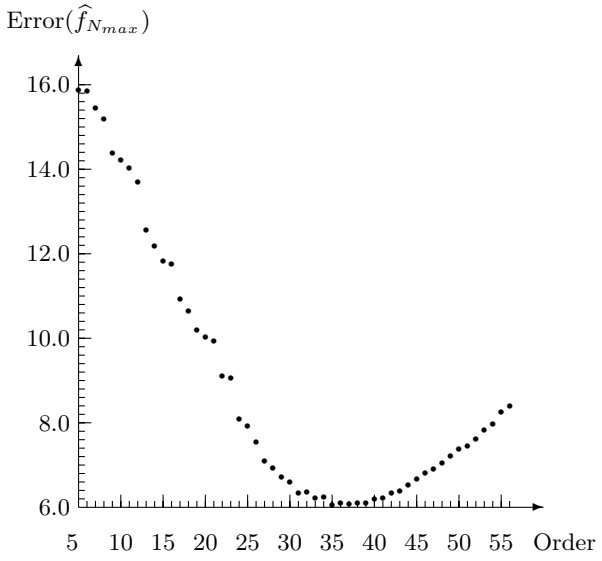


Figure 2.7: Reconstruction error Error($\widehat{f}_{N_{max}}$) versus N_{max}, $\sigma^2 = 4.0$, $\Delta^2 = 0.083\ldots$

this observation and this is illustrated in Fig. 2.9. The same Chinese character C_1 and the noise model shown in (2.33) are employed. The result is averaged over 10 runs and the noise varies from $\sigma^2 = 4.0$ to $\sigma^2 = 25.0$. It is interesting to note that N^*_{max} can be unaltered over some intervals of σ^2.

2.7. Data-Driven Selection of Number of Moments

The aforementioned results clearly show that there exists an optimal number of moments N^*_{max} which minimizes the reconstruction error Error($\widehat{f}_{N_{max}}$). This number, however, depends on the unknown $f(x, y)$ and the usally unknown σ^2.

A simple heuristics for selecting N_{max} could rely on displaying images $\widehat{f}_{N_{max}}$ for $N_{max} = 0, 1, 2, \ldots$ and assessing their quality by some subjective criteria. It is worth noting that our reconstruction algorithm in (2.28) has a very simple structure for its updating from $N_{max} - 1$ to N_{max}.

In fact by virtue of (2.28), we have

$$\widehat{f}_{N_{max}}(x, y) = \widehat{f}_{N_{max}-1}(x, y) + \sum_{s=0}^{N_{max}} \tau_{N_{max}-s,s} \widehat{\lambda}_{N_{max}-s,s} P_{N_{max}-s}(x) P_s(y). \quad (2.48)$$

Although this approach can be satisfied in some cases, it is more important to

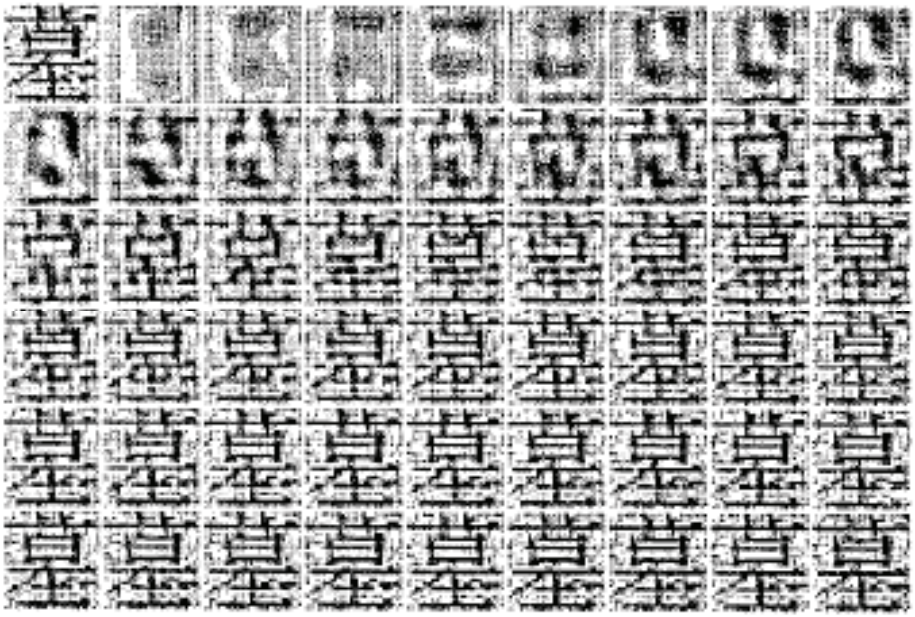


Figure 2.8: Noisy version of the character C_1, with $\sigma^2 = 4.0$, and its reconstructed versions from Legendre moments of orders 4–56

propose a rigorous and automatic technique which can estimate N^*_{max} directly from the available data.

Such a problem can be tackled by using the well-established in statistics cross-validation methodology [41, 92], see [90, 173, 223] for a few studies of these techniques in computer vision problems. We will discuss a number of prescriptions for automatic selection of N_{max}.

It is clear that $\text{Error}(\widehat{f}_{N_{max}})$ can be written as

$$\iint_{\Omega} \widehat{f}^{\,2}_{N_{max}}(x,y)dxdy - 2\iint_{\Omega} \widehat{f}_{N_{max}}(x,y)f(x,y)dxdy + \iint_{\Omega} f^2(x,y)dxdy. \quad (2.49)$$

Since the last term is independent of N_{max} and by virtue of Parseval's formula the minimization of $\text{Error}(\widehat{f}_{N_{max}})$ is equivalent to taking the minimizer of the following criterion

$$S(N_{max}) = \sum_{p=0}^{N_{max}} \sum_{q=0}^{p} \tau_{p-q,q} \widehat{\lambda}^2_{p-q,q} - 2 \sum_{p=0}^{N_{max}} \sum_{p=0}^{q} \tau_{p-q,q} \widehat{\lambda}_{p-q,q} \lambda_{p-q,q}. \quad (2.50)$$

Table 2.3: Square reconstruction error Error($\widehat{f}_{N_{max}}$) = $I(N)$ with $\sigma^2 = 4.0$

Order	$I(N)$	Order	$I(N)$	Order	$I(N)$
3	1.628e+01	21	9.928e+00	39	6.099e+00
4	1.604e+01	22	9.110e+00	40	6.193e+00
5	1.587e+01	23	9.067e+00	41	6.217e+00
6	1.586e+01	24	8.097e+00	42	6.323e+00
7	1.544e+01	25	7.924e+00	43	6.377e+00
8	1.518e+01	26	7.556e+00	44	6.529e+00
9	1.439e+01	27	7.101e+00	45	6.660e+00
10	1.421e+01	28	6.939e+00	46	6.814e+00
11	1.403e+01	29	6.718e+00	47	6.898e+00
12	1.369e+01	30	6.610e+00	48	7.046e+00
13	1.256e+01	31	6.349e+00	49	7.223e+00
14	1.219e+01	32	6.359e+00	50	7.386e+00
15	1.183e+01	33	6.219e+00	51	7.444e+00
16	1.175e+01	34	6.243e+00	52	7.606e+00
17	1.093e+01	35	6.065e+00	53	7.818e+00
18	1.065e+01	36	6.103e+00	54	7.982e+00
19	1.020e+01	37	6.069e+00	55	8.258e+00
20	1.003e+01	38	6.105e+00	56	8.407e+00

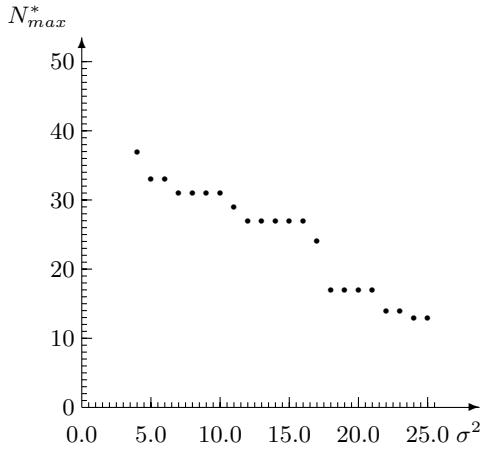


Figure 2.9: Optimal number of moments N^*_{max} versus the noise variance

The cross-validation techniques require estimating

$$\widehat{\lambda}_{pq}\lambda_{pq} = \eta_{pq},$$

say, from two non-overlapping sets of data. This can be carried out in a number of ways. For example, one could determine a simple estimate of λ_{pq}, called $\overline{\lambda}_{pq}$, based only on data coming from a small window, and discard that data when computing the main estimate of λ_{pq}, called $\widehat{\widetilde{\lambda}}_{pq}$.

This would lead to the following estimate of $S(N_{max})$

$$\widehat{S}(N_{max}) = \sum_{p=0}^{N_{max}} \sum_{q=0}^{p} \tau_{p-q,q} \widehat{\widetilde{\lambda}}^2_{p-q,q} - 2 \sum_{p=0}^{N_{max}} \sum_{q=0}^{p} \tau_{p-q,q} \widehat{\widetilde{\lambda}}_{p-q,q} \overline{\lambda}_{p-q,q}.$$

We refer to [90] for some theoretical development of this approach in the context of the image restoration problem. On the other hand in [173] the leave-one-out approach has been used yielding the following estimate of η_{pq}

$$\widehat{\eta}_{pq} = \sum_{i=1}^{n} \sum_{j=1}^{n} g(x_i, y_j) \widehat{\lambda}_{pq,ij} h_{pq}(x_i, y_j),$$

where $\widehat{\lambda}_{pq,ij}$ is the leave-one-out version of $\widehat{\lambda}_{pq}$ and it can be approximately determined as

$$\widehat{\lambda}_{pq} - \Delta^2 (\partial_x g)(x_i, y_j).$$

Here $(\partial_x g)(x, y)$ is an estimate of $\dfrac{\partial f(x, y)}{\partial x}$ obtained from the data $\{g(x_i, y_j)\}$. Some theoretical properties of this estimate have been established in [173].

Let us observe that both aforementioned techniques do not use any prior knowledge about the noise model and about the true image.

For a large class of images (Remark 2.4) the second term in the decomposition in (2.44) is of the smaller order than the first one. In such a case and due to Theorem 2.6 the minimization of $E(\mathrm{Error}(\widehat{f}_{N_{max}}))$ is equivalent to seeking a minimum of the following criterion

$$L(N_{max}) = \sigma^2 \frac{(N_{max}+1)(N_{max}+2)}{2} \Delta^2 - \sum_{p=0}^{N_{max}} \sum_{q=0}^{p} \tau_{p-q,q}\, \lambda^2_{p-q,q}. \tag{2.51}$$

An empirical version of $L(N_{max})$ requires estimation of λ^2_{pq} and σ^2. For the latter problem, see [89] and [223]. On the other hand, the term λ^2_{pq} due again to the cross-validation principle can be estimated by

$$\sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{l=1}^{n} \sum_{t=1}^{n} g(x_i, y_j)\, g(x_l, y_t)\, h_{pq}(x_i, y_j)\, h_{pq}(x_l, y_t),$$

where the summation is carried out for all $(i,j) \neq (l,t)$. Plugging this into (2.51) yields an estimate of $L(N_{max})$. No theoretical properties of such selector of N_{max} are known. It is worth noting that N_{max} minimizing (2.50) is a function of the data at hand, whereas a minimizer of $L(N_{max})$ depends on $f(x,y)$, σ^2 and Δ.

Yet another class of cross-validation techniques make use of the residual sum of squares function

$$RSS(N_{max}) = \Delta^2 \sum_{i=1}^{n} \sum_{j=1}^{n} (g(x_i, y_j) - \widehat{f}_{N_{max}}(x_i, y_j))^2.$$

Minimization of $RSS(N_{max})$ leads to an unacceptably large value of N_{max}. The so-called generalized cross-validation approach [41,230] uses the penalized version of $RSS(N_{max})$.

Hence, a criterion for selecting N_{max} is of the following form

$$GCV(N_{max}) = RSS(N_{max})\Phi(N_{max}), \tag{2.52}$$

where the penalty $\Phi(N_{max})$ can be chosen as

$$\Phi(N_{max}) = \left(1 - \gamma(\sigma^2)\frac{(N_{max}+1)(N_{max}+2)}{2}\Delta^2\right)^{-\alpha},$$

where $\alpha > 0$ and $\gamma(\sigma^2)$ is a non-decreasing function of σ^2. See [92] for a detailed discussion of such selectors in the context of nonparametric curve estimation.

2.8. Generalized Orthogonal Moments

Teague [219] has first introduced a class of orthogonal moments utilizing the theory of classical orthogonal polynomials. In particular he employed Legendre and Zernike polynomials and the corresponding moments which have been extensively used in numerous applications since then. In the previous sections we have studied the Legendre polynomials and moments which represent a special type of classical orthogonal polynomials [217]. In the subsequent section we wish to explore a larger class of classical orthogonal polynomials called Gegenbauer polynomials or ultraspherical polynomials [217]. This class is characterized by the presence of the scaling parameter which can define a larger class of orthogonal polynomials. Indeed the Legendre and other previously introduced orthogonal moments are obtained as special cases of the Gegenbauer moments.

We argue that a reconstruction formula based on the Gegenbauer polynomials and corresponding moments can achieve an improved performance by selecting the scaling factor which yields the minimum of the reconstruction error.

Furthermore, the scaling parameter may provide a trade-off between global and local image features. In fact, a large value of the parameter gives a purely local representation of an image, whereas its small values correspond to classical global moments.

The classical orthogonal moments attain a zero value of redundancy measures and, in addition, the reconstruction procedure as we have pointed out does not need regularization as it simply adds up the individual contributions of each moment. These favourable properties of the orthogonal moments have lead to a number of studies on computational, theoretical and application aspects of this type of moments [144,145,173,180,221,226]. In particular, the fundamental issues concerning the error analysis, accuracy, and noise sensitivity have been examined for the Legendre and Zernike moments [144, 145, 173, 180].

Recently a new type of orthogonal moments [158,248] generated by orthogonal polynomials of a discrete variable has been introduced. The main goal of these studies was to overcome the problem of sampling error. This is an interesting alternative assuming that a true image is discrete with a fixed degree of spatial resolution. If the resolution, however, is increased then one has to recalculate the discrete orthogonal polynomials. This is not the case with analog orthogonal functions which are able to copy with any degree of resolution. In the concluding section of this chapter we give a brief critical overview of the discrete orthogonal moments with main emphasis on the reconstruction problem from noisy images.

2.9. Gegenbauer Moments

A class of orthogonal moments examined in this chapter originates from the so-called Gegenbauer polynomials $\{G_p(x,\lambda)\}$ [217] which can be computed from the following recurrence formula

$$(p+1)G_{p+1}(x,\lambda) = 2(p+\lambda)xG_p(x,\lambda) - (p+2\lambda-1)G_{p-1}(x,\lambda) \tag{2.53}$$

for $p = 1, 2, \ldots$, where $G_0(x,\lambda) = 1$, $G_1(x,\lambda) = 2\lambda x$, $-1 \le x \le 1$ and $\lambda > -0.5$ is the scaling factor. The polynomials $\{G_p(x,\lambda)\}$ are orthogonal with respect to the weight function $w(x) = (1-x^2)^{\lambda-1/2}$, i.e.,

$$\int_{-1}^{1} w(x)G_p(x,\lambda)G_q(x,\lambda)dx = \delta_{p,q}C_p(\lambda), \tag{2.54}$$

where the normalized constant $C_p(\lambda)$ is given by

$$C_p(\lambda) = \sqrt{\pi}\frac{\Gamma(\lambda+1/2)}{\Gamma(\lambda)(p+\lambda)}G_p(1,\lambda).$$

The scaling factor $\lambda > -0.5$ can define a large class of orthogonal polynomials. In fact, $G_p(x, 0.5)$ is the Legendre polynomial of the pth order and this case was considered in the previous sections. The case $\{G_p(x, 1\})$ defines the class of Chebyshev polynomials of the second kind [217]. The case $G_p(x, 0)$ should be understood as follows

$$G_p(x, 0) = \lim_{\lambda \to 0} \frac{1}{\lambda} G_p(x, \lambda) = \frac{2}{p} T_p(x), \quad p \geq 1, \tag{2.55}$$

with $G_0(x, 0) = 1$, where $T_p(x)$ are the Chebyshev polynomials of the first kind. Yet a different limiting case occurs if $\lambda \to \infty$. Then it can be shown [217] that

$$\lim_{\lambda \to \infty} \frac{G_p(x, \lambda)}{G_p(1, \lambda)} = x^p. \tag{2.56}$$

This formula explains the scaling role played by λ as it gives a local representation of a function by the Gegenbauer moments, see the discussion on this issue at the end of this section.

The orthogonality and completeness of $\{G_p(x, \lambda)\}$ allows us to represent an image via a truncated Gegenbauer series. First let us note that $G_p(x, \lambda)G_q(y, \lambda)$ defines the 2-D Gegenbauer polynomial of order $p + q$. Fig. 2.10 depicts density plots of $G_p(x, \lambda)G_q(y, \lambda)$ for a few combinations of (p, q) with $\lambda = 0.7$.

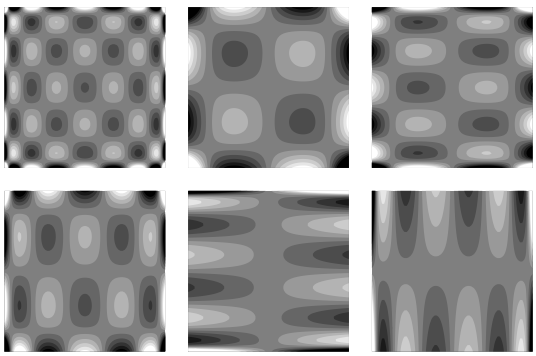


Figure 2.10: Gegenbauer polynomials $G_p(x, 0.7)G_q(y, 0.7)$ for $(p, q) \in \{(8, 6), (3, 3), (3, 6); (6, 3), (1, 7), (7, 1)\}$

Using the fact that $\{G_p(x, \lambda)G_q(y, \lambda)\}$ forms an orthogonal and basis on $\Omega = [-1, 1]^2$ one can represent an image function $f(x, y)$ in the following way

$$f_N(x, y) = \sum_{p=0}^{N} \sum_{q=0}^{p} \tau_{p-q,q} A_{p-q,q}(\lambda) G_{p-q}(x, \lambda) G_q(y, \lambda), \tag{2.57}$$

where $\tau_{pq} = (C_p(\lambda)C_q(\lambda))^{-1}$ is the normalizing factor and A_{pq} is the (p, q) Gegenbauer moment of $f(x, y)$ which is defined by

$$A_{pq}(\lambda) = \int_{-1}^{1}\int_{-1}^{1} f(x, y)G_p(x, \lambda)G_q(y, \lambda)w(x)w(y)dxdy. \tag{2.58}$$

It should be expected that the representation in (2.57) can converge to the true image by making N sufficiently large. If, however, we define an upper limit to N, then it is less obvious that the approximation error can be minimized by an optimum choice of the scale factor λ. In fact for small and moderate values of λ we have a global representation of $f(x, y)$ in terms of the Gegenbauer moments defined in (2.58). On the other hand if $\lambda \to \infty$ then by (2.56) and by some tedious but straightforward algebra we can show that the (p, q) term in (2.57) tends to

$$\frac{1}{p!}\left(x\frac{\partial}{\partial x} + y\frac{\partial}{\partial y}\right)^p f(0, 0), \tag{2.59}$$

being the pth term in the Taylor's expansion of $f(x, y)$ about the point $(0, 0)$. Here it has been assumed that $f(x, y)$ is smooth, i.e., it has derivatives up to the order N. As a consequence of this result we can obtain that $\lim_{\lambda\to\infty} f_N(x - x_0, y - y_0)$ is the Nth order Taylor's expansion of $f(x, y)$ about the point (x_0, y_0). Hence for large λ we obtain a purely local approximation of the image function. This intuitive explanation seems to indicate that there is λ yielding the best possible representation of a given image by the Gegenbauer moments. We demonstrate experimentally in this section that this is really the case.

2.10. Image Reconstruction from Gegenbauer Moments

In practice one can only observe a digital and noisy version of the image function and has to calculate corresponding moments. Hence let $\{g(x_i, y_j), 1 \le i, j \le n\}$ be a noisy version of $f(x, y)$ observed over the $n \times n$ array of pixels with the sampling period Δ. Here the noise model for obtaining $g(x_i, y_j)$ is the same as that examined in Section 2.5, i.e., $g(x_i, y_j) = f(x_i, y_j) + z(x_i, y_j)$, see (2.33).

Then analogously as in Section 2.5 we can define the following estimate of $A_{pq}(\lambda)$

$$\widehat{A}_{pq}(\lambda) = \sum_{i=1}^{n}\sum_{j=1}^{n} h_{p,q}(x_i, y_j)g(x_i, y_j), \tag{2.60}$$

where

$$h_{p,q}(x_i, y_j) = \int_{x_i-\Delta/2}^{x_i+\Delta/2} \int_{y_j-\Delta/2}^{y_j+\Delta/2} G_p(x,\lambda) G_q(y,\lambda) w(x) w(y) dxdy \tag{2.61}$$

represents the weighted integration of $G_p(x,\lambda)G_q(y,\lambda)$ over the (i,j) pixel. The integral appearing in (2.61) can be evaluated by high precision numerical integration formulas mentioned in Section 2.3. The simplest strategy to estimate the integral is the one point formula $\Delta^2 G_p(x_i,\lambda)G_q(y_j,\lambda)w(x_i)w(y_j)$.

The performance of the moment descriptors $\widehat{A}_{pq}(\lambda)$ is best assessed by means of the image reconstruction. Thus plugging (2.61) into (2.57) yields our reconstruction formula from the first N Gegenbauer moments

$$\widehat{f}_N(x,y) = \sum_{p=0}^{N} \sum_{q=0}^{p} \tau_{p-q,q} \widehat{A}_{p-q,q}(\lambda) G_{p-q}(x,\lambda) G_q(y,\lambda). \tag{2.62}$$

The mean integrated-squared error of $\widehat{f}_N(x,y) - f(x,y)$ can serve as a natural measure of assessing the performance of our reconstruction formula, i.e.,

$$Err(N,\lambda) = E \int_{-1}^{1} \int_{-1}^{1} (\widehat{f}_N(x,y) - f(x,y))^2 w(x) w(y) dxdy. \tag{2.63}$$

Due to Parseval's identity the error $Err(N,\lambda)$ can be decomposed into three unrelated terms, i.e., the noise error

$$Err^Z(N,\lambda) = \sum_{p=0}^{N} \sum_{q=0}^{p} \tau_{p-q,q} \operatorname{var}(\widehat{A}_{p-q,q}(\lambda)),$$

the discretization error

$$Err^D(N,\lambda) = \sum_{p=0}^{N} \sum_{q=0}^{p} \tau_{p-q,q} (E\widehat{A}_{p-q,q}(\lambda) - A_{p-q,q}(\lambda))^2,$$

and the approximation error

$$Err^A(N,\lambda) = \sum_{p=N+1}^{\infty} \sum_{q=0}^{p} \tau_{p-q,q} A^2_{p-q,q}(\lambda).$$

Arguing as in Section 2.6 it can be easily shown that both the noise $Err^Z(N,\lambda)$ and discretization errors $Err^D(N,\lambda)$ are of order $\Delta^2(N+1)(N+2)/2$ provided that the image function is in the class $\mathrm{BV}(\Omega)$. These upper bounds are uniform with respect to the scaling parameter λ.

On the other hand the approximation error $Err^A(N,\lambda)$ decreases with N and is a complicated function of λ. Nevertheless, under some mild regularity conditions on $f(x,y)$ it is conjectured that there is λ which minimizes $Err^A(N,\lambda)$. Clearly such an optimal λ depends not only on $f(x,y)$ but also on the number of moments N used for reconstruction. It is still an open problem of how to evaluate optimal λ for a large class of image functions and arbitrary values of N. The fact that the approximation error can be optimized with respect to the scaling parameter λ is due to the amount of local information contributing to the overall approximation error. In fact, we have pointed out that for large λ we may obtain local moment descriptors, whereas small λ is useful for the global, Legendre type, moments. This observation could lead to an interesting localized version of moments being sensitive to boundaries. Accordingly, we may compute the moments of a differentiated image (like the magnitude of the image gradient) instead of the original image. This could be compared to differential invariants used for recognition of a planar curve regardless of the point of view from which the curve is seen [234].

Analogously as in Section 2.6 we can show that the approximation error $Err^A(N,\lambda)$ as a function of N decays at least as fast as $O(N^{-1})$ for all images with bounded variation. This gives an optimal choice of N of order $N^\star = \alpha\Delta^{-2/3}$ yielding $Err(N^\star,\lambda) = O(\Delta^{2/3})$. Again these are upper bounds and the dependence on λ has been ignored. Nevertheless, the aforementioned discussion reveals that large values of λ yield localized moments which can be sensitive to noise. This may lead to another trade-off in selecting an optimal λ.

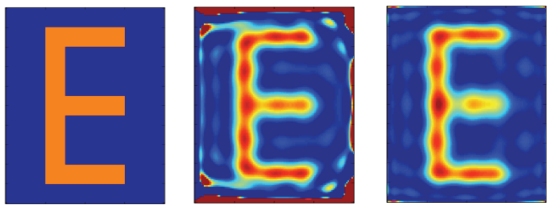


Figure 2.11: Image reconstruction of letter "E" with $N = 16$ and $\lambda = 4, \lambda = 0.7$

To illustrate the above discussion let us consider a simple numerical problem of reconstruction of the letters "E" and "A" on a 60×60 pixel grid, i.e., when $\Delta = 1/30$. The maximal number of the moments taken is $N = 16$. Fig. 2.11 shows the reconstruction of the letter "E" for $\lambda = 4, \lambda = 0.7$ by the estimate $\widehat{f}_N(x,y)$. Fig. 2.12 depicts the reconstruction of the letter "A" for $\lambda = 0, \lambda = 0.7$. In both examples we have observed noise free pictures. The reconstruction error $Err(16,\lambda)$ versus λ is displayed in Fig. 2.13. It is apparent that there is λ

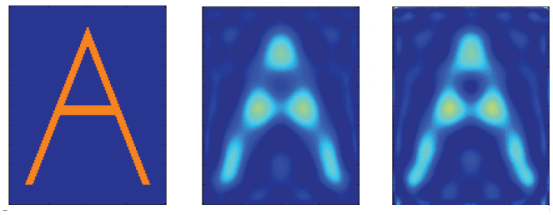


Figure 2.12: Image reconstruction of letter "A" with $N = 16$ and $\lambda = 0, \lambda = 0.7$

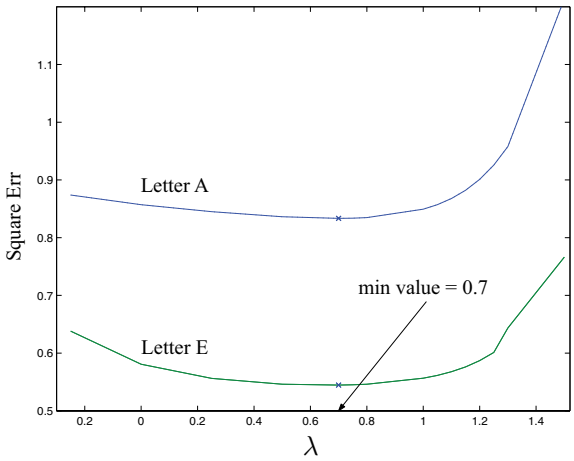


Figure 2.13: The reconstruction error for letters "E", "A" versus λ, with $N = 16$

minimizing the error. Surprisingly the optimal $\lambda^\star = 0.7$ is the same for both patterns under examination. Experiments were performed for reconstruction of other letters giving the optimal $\lambda^\star$ in the range $[0.6, 0.8]$. Recall that $\lambda = 0.5$ corresponds to the Legendre moments.

2.11. Discrete Orthogonal Moments

Thus far we have considered the image model assuming that a true image function $f(x, y)$ is analog and belongs the class of functions of bounded variation. This assumption plays an important role in our theoretical developments and is also crucial in the current mathematical theory of image analysis. Consequently our asymptotic analysis of the moment accuracy and reconstruction power is modeled on grids with increasing resolution. There have been recent studies [158,248]

on moment descriptors utilizing orthogonal polynomials on a finite set of points. Hence an image model is an array of $n \times n$ pixels defined, without loss of generality, on the set $\Omega = \{(i,j), 0 \leq i,j \leq n\}$. Thus we have the digital image $\{f(x,y), (x,y) \in \Omega\}$.

The orthogonal system on Ω is defined as $\{P_p(x)P_q(y), 0 \leq p,q \leq n\}$ for $(x,y) \in \Omega$, where the orthogonality of $\{P_p(x), 0 \leq p \leq n\}$ is defined as follows

$$\sum_{x=0}^{n} P_p(x)P_q(x) = C_p(n)\delta_{pq}.$$

Here $C_p(n) = \sum_{x=0}^{n} P_p^2(x)$ is the normalizing constant.

Consequently we can represent the digital image $\{f(x,y), 0 \leq x,y \leq n\}$ by the following expansion employing the first $(T+1)^2$ discrete moments

$$f_T(x,y) = \sum_{p=0}^{T}\sum_{q=0}^{T} \tau_{pq}\lambda_{pq}P_p(x)P_q(y), \quad (x,y) \in \Omega, \tag{2.64}$$

where $\tau_{pq} = (C_p(n)C_q(n))^{-1}$ is the normalizing constant and

$$\lambda_{pq} = \sum_{i=0}^{n}\sum_{j=0}^{n} f(i,j)P_p(i)P_q(j)$$

defines the (p,q) order discrete moment with respect to the basis $\{P_p(x)P_q(y), 0 \leq p,q \leq n\}$.

A concrete example of $\{P_p(x), 1 \leq p \leq n\}$ is a discrete analog of Legendre orthogonal polynomials

$$P_p(x) = \sum_{s=0}^{p} \frac{(-1)^s \binom{p}{s}\binom{p+s}{s}}{n(n-1)\dots(n-s+1)} x^{[s]}$$

for $0 \leq p \leq n$ and $0 \leq x \leq n$, where $x^{[s]} = x(x-1)\dots(x-s+1)$ with $x^{[0]} = 1$. In particular, $P_0(x) = 1$, $P_1(x) = 1-2x/n$, $P_2(x) = 1-6x/n+6x(x-1)/(n(n-1))$.

Computational and experimental studies of the discrete orthogonal moments for noise free images have been worked out in [158,248]. No reconstruction aspects of the proposed moments have been given.

It is clear that for the reconstruction formula in (2.64) the squared error

$$Err(f_T) = \sum_{x=0}^{n}\sum_{y=0}^{n}(f_T(x,y) - f(x,y))^2$$

does not have any discretization component.

Nevertheless, if we apply $f_T(x, y)$ for noisy image (with $f(i, j)$ replaced by $g(i, j) = f(i, j) + z(i, j)$) then using the techniques developed in this section we can show that the variance term of the error is $\sigma^2(T+1)^2$. This is an unacceptably large value and reveals that the algorithm $f_T(x, y)$ cannot behave well for noisy images. One way of overcoming this difficulty is to smooth noisy data $\{g(i, j), 0 \leq i, j \leq n\}$ before they are applied in the reconstruction process. This leads to a modified class of discrete moments deserving further studies. It is also worth mentioning that the issue of obtaining invariant features from discrete orthogonal moments remains open.

2.12. Conclusions

In this chapter, we have developed the basic methodology for the problems of accuracy and efficiency in moment computing. Both geometric and orthogonal moments were examined although the main emphasis was given to the latter class. Based on the improved moment computing techniques, image reconstruction algorithms based on the Legendre and Gegenbauer moments from discrete and noise data were proposed. The delicate problem of selecting an optimal number of moments from the available data has been examined. A discussion of this issue based on the cross-validation methodology has been provided. In summary the following main results have been obtained.

- The numerical techniques for high quality calculation of geometric and orthogonal moments have been proposed.
- Bounds on the precision of the proposed estimates have been established.
- The algorithms for image recovery from Legendre moments have been proposed and their asymptotic analysis has been investigated.
- Numerous simulation studies have been carried out to confirm our basic theory.
- The estimates of geometric and orthogonal moments from noisy images have been defined and their statistical accuracy has been evaluated.
- The rate of convergence of the Legendre moment based image estimate has been evaluated.
- Various algorithms for automatic selection of optimal number of moments have been derived.
- The aforementioned theory has been further extended to a general class of orthogonal moments and moments calculated in the digital domain.

2.13. Appendix

Proof of Theorem 2.1

Let

$$A_{ij} = \left[x_i - \frac{\Delta}{2}, x_i + \frac{\Delta}{2}\right] \times \left[y_j - \frac{\Delta}{2}, y_j + \frac{\Delta}{2}\right]$$

denote the (i, j) pixel.

First let us observe that

$$m_{pq} = \sum_{i=1}^{n}\sum_{j=1}^{n} \iint_{A_{ij}} f(x, y)x^p y^q dx\, dy.$$

By this and the Cauchy-Schwartz inequality, we obtain

$$|\widehat{m}_{pq} - m_{pq}| \leq 2((2p+1)(2q+1))^{-\frac{1}{2}} \cdot \left(\sum_{i=1}^{n}\sum_{j=1}^{n} \iint_{A_{ij}} (f(x_i, y_j) - f(x, y))^2 dx\, dy\right)^{\frac{1}{2}}.$$

The term in brackets does not exceed

$$\sum_{i=1}^{n}\sum_{j=1}^{n} Osc_{ij}(f) \iint_{A_{ij}} |f(x_i, y_j) - f(x, y)| dxdy \leq f_{max}\, \mathrm{TV}(f)\Delta^2,$$

where

$$Osc_{ij}(f) = \sup\{|f(w_1, w_2) - f(z_1, z_2)| : (w_1, w_2), (z_1, z_2) \in A_{ij}\}$$

is the oscillation of $f(x, y)$ over the pixel A_{ij}. Since $\sum_{i=1}^{n}\sum_{j=1}^{n} Osc_{ij}(f) = \mathrm{TV}(f)$ the proof of Theorem 2.1 has been completed. □

Chapter 3

Image Analysis by Orthogonal Radial Moments

In this chapter we consider the problem of image characterization, representation, and reconstruction utilizing a class of orthogonal and complete polynomials defined over a unit disk often referred to as the Zernike polynomials. This is carried out by projecting the input image onto the space spanned by the Zernike polynomials yielding the concept of Zernike moments. This class of moments has a distinctive property of being invariant to rotations and reflections and thus it has found a number of important applications in pattern recognition, image analysis, ophthalmology, medical imaging, optical engineering, and watermarking. Furthermore, this orthogonal set defines the MPEG-7 standard for visual shape descriptors. In this chapter an introduction to computational and reconstruction aspects of Zernike moments is given. We illustrate our discussion by numerous numerical studies.

In summary, the main contributions of this chapter are the following:

- Basic analytical properties of orthogonal radial polynomials being invariant in form are reviewed.
- Invariant properties of Zernike moments are derived. This includes the invariance for rotations and reflections.
- An improved technique for high quality calculation of Zernike moments is proposed.
- The error analysis of the numerical estimate of the Zernike moment is carried out.

- A connection between the accuracy problem and the analytic number theory of lattice points is established.
- High order numerical integration methods for Zernike moments computing are developed.
- Bounds for the precision of given estimates are established.
- Modified Zernike moments with the reduced geometric error are proposed.
- Numerical studies are carried out to confirm the basic theory.

3.1. Introduction

Teague [219] has suggested the use of orthogonal moments in image analysis based on a certain class of orthogonal polynomials to overcome the aforementioned shortcomings associated with the geometric moments. In particular, he has proposed to use the orthogonal moments defined in terms of the Legendre and Zernike polynomials. A number of studies on the use of the Legendre and Zernike moments in shape analysis and classification have been carried out [8, 121, 122, 199, 219, 221, 226]. Nevertheless, the error analysis and analytic characterization of the orthogonal moments have been rarely investigated. In the previous chapter the vulnerability of the orthogonal Legendre moments against discretization and noise have been examined, see also [173] and [144]. The Legendre moments, however, despite their good reconstruction properties, are not invariant to rotations. The Zernike orthogonal moments, being invariant to rotations and reflections, have been suggested to be an attractive alternative [219]. This set of orthogonal functions has been introduced by Zernike [250] as a basic tool for representation of a wavefront function for optical systems with circular pupils. Since then the radial polynomials have been found important in applications ranging from pattern recognition, shape analysis, optical engineering, medical imaging to eye diagnostic [18,33,102,107–109]. Recently, under the name of an angular radial transform, moments based on radial orthogonal polynomials have been adopted in MPEG-7 as a basic region-based shape descriptor [17].

Furthermore, in numerous empirical studies on radial orthogonal moments, like Zernike, it has been found that they are the best choices for object representation, see in particular [226] for such recommendations in the context of character recognition, and [251] in the field of shape retrieval.

Radial orthogonal moments are defined on the unit circle, i.e., the image plane is $D = \{(x, y) : x^2 + y^2 \leq 1\}$. This geometry of image plane is natural for a number of natural images including the analysis of data from positron scanning devices [33, 115, 157], the diffraction theory of optical aberrations [18, 102, 250],

invariant pattern recognition [1, 101, 121, 173, 228, 232], image analysis [124, 127, 144, 145, 173, 199, 219, 221], and statistical models of circular data [22]. In Fig. 3.1 we depict some examples of circular type images.

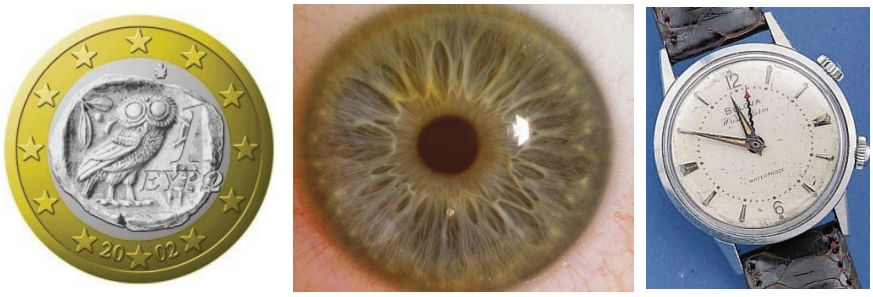


Figure 3.1: Examples of natural images of the circular form

In this chapter, we introduce the problem of computational and reconstruction properties of radial moments. First of all we give formulas for computing the radial moments from discrete data and describe the main sources of errors in their evaluation. In particular, the numerical accuracy of the radial moments computing is examined. In this respect, some new techniques utilizing various two-dimensional numerical integration methods for increasing the overall accuracy of the radial moments are proposed. We illustrate our findings by showing the reconstruction performance of the radial moments. As has been indicated in the previous chapter our main emphasis is on the problem of accuracy. Fast algorithms for calculating the radial moments have been proposed elsewhere, see [159].

3.2. Orthogonal Radial Polynomials

In order to define a class of radial orthogonal moments, we need to introduce the concept of corresponding orthogonal radial functions. Such a set of functions, denoted by $V(x, y)$, is defined on the unit disk $D = \{(x, y) : x^2 + y^2 \leq 1\}$ and assumed to be invariant with respect to any rotation

$$\begin{aligned} x' &= x\cos\theta + y\sin\theta, \\ y' &= -x\sin\theta + y\cos\theta. \end{aligned}$$

The invariance is meant as the property that when the rotation is applied, each element of the set is transformed into an element of the same form, i.e.,

$$V(x, y) = A(\theta)V(x', y'), \tag{3.1}$$

where $A(\theta)$ is a continuous function of θ with period 2π in θ and satisfies the initial condition $A(0) = 1$. The function $V(x, y)$ which meets (3.1) is said to be invariant in form. The following fundamental result concerning invariant polynomials was proved in [13].

Theorem 3.1. *A polynomial $V(x, y)$ of degree p will be invariant in form if and only if when expressed in polar coordinates (ρ, θ) is of the form*

$$V_{pq}(\rho\cos\theta, \rho\sin\theta) = R_{pq}(\rho)e^{jq\theta},$$

where q is an integer being positive, negative, or zero and $R_{pq}(\rho)$ is a polynomial in ρ of degree p, containing no power of ρ lower than $|q|$. Moreover, $R_{pq}(\rho)$ is an even or odd polynomial according to whether q is even or odd.

It is clear that there are infinitely many sets of invariant polynomials which meet the condition of Theorem 3.1. For example, we can obtain such functions by the orthogonalization procedure of the linearly independent functions

$$\rho^{|q|}e^{jq\theta}, \rho^{|q|+\gamma}e^{jq\theta}, \rho^{|q|+2\gamma}e^{jq\theta}, ..., \tag{3.2}$$

with respect to a weight function $w(\rho), 0 \leq \rho \leq 1$ and where $0 < \gamma$.

To narrow the class of admissible invariant functions characterized by Theorem 3.1 we choose the circle polynomial $R_{pq}(\rho)$ obtained by orthogonalization procedure with $\gamma = 2$ and $w(\rho) = 1$. This yields only one set of invariant polynomials called the Zernike functions. We shall use the Zernike functions due to their unique properties and importance in applications. Indeed it was proved in [13] that the following conditions uniquely characterize the Zernike invariant functions, i.e., there is one and only one set which is:

- orthogonal on the unit disk,
- contains only such polynomials which are invariant in form,
- contains a polynomial for each permissible pair of values of p (degree) and q (angular dependence), i.e., for integer values of p and q such that $p \geq 0$, $q \geq 0, q < 0$, $p \geq |q|$, and $p - |q|$ is even.

Hence let the (p, q) order Zernike function be defined as follows

$$V_{pq}(x, y) = R_{pq}(\rho)\exp(jq\theta), \qquad x^2 + y^2 \leq 1, \tag{3.3}$$

where $\rho = \sqrt{x^2 + y^2}$ is the length of the vector from the origin to the pixel (x, y), and $\theta = \arctan(y/x)$ is the angle between the vector and the x axis. In (3.3),

$R_{pq}(\rho)$ is a polynomial in ρ of degree $p \geq 0$, containing no power of ρ lower than $|q|$. The degree p and the angular dependence q satisfy the above constraints.

It is worth noting that the case $\gamma = 1$ and $w(\rho) = 1$ in (3.2) gives the so-called pseudo-Zernike invariant functions. For this class of invariant functions the set of permissible pairs (p, q) must be such that $p \geq |q|$ and the corresponding radial function $Z_{pq}(\rho)$ is related to the Zernike radial function $R_{pq}(\rho)$ by the identity $\rho Z_{pq}(\rho^2) = R_{2p+1,2q+1}(\rho)$ [13].

The Zernike functions besides being invariant with respect to rotations have the following properties which we will find useful in our considerations.

1. They are orthogonal on the unit disk, i.e., the orthogonality relation for $\{V_{pq}(x, y)\}$ is

$$\iint_D V^*_{pq}(x, y) V_{p'q'}(x, y) dx dy = \frac{\pi}{p+1} \delta_{pp'} \delta_{qq'} \tag{3.4}$$

   or in polar coordinates

$$\int_0^{2\pi} \int_0^1 V^*_{pq}(\rho, \theta) V_{p'q'}(\rho, \theta) \rho d\rho d\theta = \frac{\pi}{p+1} \delta_{pp'} \delta_{qq'}, \tag{3.5}$$

   where the asterisk denotes the complex conjugate.

   The even and odd versions of $V_{pq}(x, y)$ are $R_{pq} \sin(q\theta)$ and $R_{pq} \cos(q\theta)$, respectively.

2. Using the above property we obtain the orthogonality relation for the radial polynomials $\{R_{pq}(\rho)\}$

$$\int_0^1 R_{pq}(\rho) R_{p'q}(\rho) \rho d\rho = \frac{1}{2(p+1)} \delta_{pp'}. \tag{3.6}$$

3. An explicit expression for the radial Zernike polynomial $R_{pq}(\rho)$ is given by

$$R_{pq}(\rho) = \sum_{l=0}^{(p-|q|)/2} (-1)^l \frac{(p-l)!}{l! \left(\frac{p+|q|}{2} - l\right)! \left(\frac{p-|q|}{2} - l\right)!} \rho^{p-2l}, \tag{3.7}$$

   with $R_{p,-q}(\rho) = R_{pq}(\rho)$ and $R_{pq}(-\rho) = (-1)^q R_{pq}(\rho)$.

4. The polynomial $R_{pq}(\rho)$ is closely related to the classical Jacobi polynomials [13, 33, 126, 217] according to the relation

$$R_{pq}(\rho) = \rho^{|q|} P_s^{(0,|q|)}(2\rho^2 - 1), \tag{3.8}$$

where $s = (p-|q|)/2$ and $P_n^{(\alpha,\beta)}(x)$ is the Jacobi polynomial of order n with the parameter (α, β). The Jacobi polynomials $\{P_n^{(\alpha,\beta)}(x)\}$ are orthogonal on the interval $[-1, 1]$ with respect to the weight function $(1-x)^\alpha(1+x)^\beta, \alpha > -1, \beta > -1$, [217].

The identity in (3.8) implies that if $p = |q|$ then

$$R_{pp}(\rho) = \rho^p, \tag{3.9}$$

while for $q = 0$ we have

$$R_{p0}(\rho) = P_{p/2}(2\rho^2 - 1), \tag{3.10}$$

where $P_s(x)$ is the s-th order classical Legendre polynomial defined on $[-1, 1]$, [206, 217].

5. The Jacobi polynomials, being in the class of classical polynomials, obey a three-term recurrence formula. Using this fact and (3.8) we can write the following recurrence relation for $R_{pq}(\rho)$

$$A_1 R_{p+2,q}(\rho) = (A_2\rho^2 + A_3)R_{pq}(\rho) + A_4 R_{p-2,q}(\rho), \tag{3.11}$$

with $R_{q-2,q}(\rho) = 0$, $R_{qq}(\rho) = \rho^q$, and where

$$\begin{aligned}
A_1 &= 2p\left(\frac{p+q}{2}+1\right)\left(\frac{p-q}{2}+1\right),\\
A_2 &= 2p(p+1)(p+2),\\
A_3 &= -q^2(p+1) - p(p+1)(p+2),\\
A_4 &= -2\left(\frac{p+q}{2}\right)\left(\frac{p-q}{2}\right)(p+2).
\end{aligned}$$

In Fig. 3.2 we depict density plots of the Zernike functions $V_{pq}(x, y)$, i.e., we plot $V_{20}(x, y)$, $Im\ V_{22}(x, y)$, $Re\ V_{31}(x, y)$, $V_{40}(x, y)$, $Re\ V_{42}(x, y)$, $Re\ V_{84}(x, y)$, where $Re\ z$, $Im\ z$ are the real and imaginary parts of the complex number z. In Fig. 3.3 we depict two $R_{pq}(\rho)$ polynomials represented by three-dimensional plots.

Remark 3.1. *The following functions*

$$F_{nm}(x, y) = P_{n-m}^{(m+1/2,m+1/2)}(x) P_m\left(\frac{y}{\sqrt{1-x^2}}\right)(1-x^2)^{m/2},$$
$$m = 0, 1, ..., n; n = 0, 1, 2, ...$$

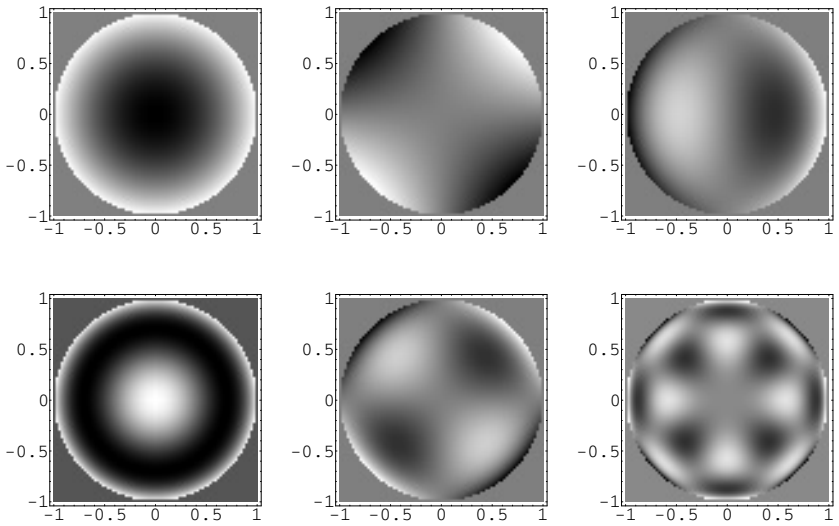
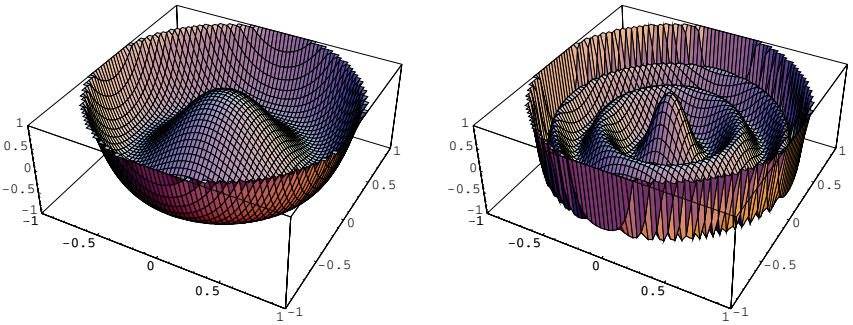


Figure 3.2: Zernike functions
$\{V_{20}(x,y), ImV_{22}(x,y), ReV_{31}(x,y); V_{40}(x,y), ReV_{42}(x,y), ReV_{84}(x,y)\}$



(a) **(b)**

Figure 3.3: Radial polynomials $R_{pq}(x,y)$ for (a) $p=4, q=0$, and (b) $p=12, q=0$.

define the orthonormal and complete system in $L_2(D)$, see [126] and references cited therein. Observe, however, that this set is not in the form of (3.3) and thus is not invariant. In [228] a class of orthogonal functions on D referred to as disk-harmonics has been proposed. They are eigenfunctions of the Laplacian operator confined to D. This class is of the invariant form as in Theorem 3.1 with $R_{pq}(\rho) = J_q(2\pi l_{pq}\rho)$, where $J_q(t)$ is the q-th order Bessel function of the first kind. Here l_{pq} is a sequence of numbers that must be selected appropriately in order to assure orthogonality and completeness. We refer to [116, 228] and

references cited therein for some theory and applications of such orthogonal bases. Observe, however, that this set has no simple connections to classical orthogonal systems and therefore its analytical properties are difficult to establish.

Finally let us also mention another interesting property of the Zernike functions. It has been shown in [211] that the generalization of the concept of prolate functions to the circular domain leads to a class of functions being closely related to the Zernike radial polynomials $R_{pq}(\rho)$.

3.3. Zernike Moments

The completeness and orthogonality of $\{V_{pq}(x,y)\}$ allow us to represent any square integrable image function $f(x,y)$ defined on the unit disk in the following series

$$f(x,y) = \sum_{p=0}^{\infty} \sum_{q=-p}^{p} \tau_p A_{pq} V_{pq}(x,y), \qquad p - |q| = \text{even}, \tag{3.12}$$

where, due to (3.4), $\tau_p = (p+1)/\pi$ is the normalizing constant and A_{pq} is the Zernike moment of order p with repetition q, i.e.,

$$A_{pq} = \iint_D f(x,y) V_{pq}^*(x,y) dx dy. \tag{3.13}$$

Recalling (3.3), the polar coordinates version of A_{pq} takes the following form

$$A_{pq} = \int_0^{2\pi} \int_0^1 \widetilde{f}(\rho,\theta) R_{pq}(\rho) \exp(-jq\theta) \rho d\rho d\theta. \tag{3.14}$$

where $\widetilde{f}(\rho,\theta) = f(\rho\cos\theta, \rho\sin\theta)$. The above representation for the Zernike moments allows us to infer about their behavior under different symmetry conditions on $f(x,y)$. Throughout the whole book we assume that the origin of the coordinate system is the center of image symmetry. Hence if $f(x,y)$ is symmetrical for every direction θ, i.e., $\widetilde{f}(\rho,\theta)$ is an even function of θ, then we have

$$A_{pq} = \int_0^{2\pi} \int_0^1 \widetilde{f}(\rho,\theta) \cos(q\theta) R_{pq}(\rho) \rho d\rho d\theta.$$

Hence in this case A_{pq} is a real number.

The fundamental property of the Zernike functions and consequently coefficients $\{A_{pq}\}$ is their rotational invariance, i.e., the rotational invariance of

$\{|A_{pq}|\}$ takes place. Hence if $f(x,y)$ is rotated through the angle α, then using (3.14) we can obtain that the Zernike coefficient A^r_{pq} of the rotated image $\widetilde{f}^r(\rho,\theta) = \widetilde{f}(\rho,\theta-\alpha)$ is given by

$$A^r_{pq} = e^{-jq\alpha}A_{pq}. \tag{3.15}$$

This identity explains the role played by the angular repetition q. Hence if $q=2$ then A_{pq} is repeated twice while the image is rotated through 360°.

Similarly for the reflected function $f(-x,-y)$ the Zernike moment is equal to $(-1)^q A_{pq}$. Generally, when the image is reflected across a line rotated through the angle β the Zernike moment changes in the following way

$$A^{refl}_{pq} = e^{-j2q\beta}A^*_{pq}, \tag{3.16}$$

where z^* is the conjugate of z.

Thus, the magnitudes of $\{A_{pq}\}$ can be used as invariant features with respect to all rotations and reflections. In fact it is common in object recognition to use $\{|A_{pq}|, p \geq |q|\}$ for a few values of (p,q) as a feature vector. See [1, 121, 124, 145, 199, 219, 231] and the references cited therein for applications of Zernike coefficients in various aspects of pattern recognition. Let us also mention that formula (3.15) can be used to estimate the rotation angle α [124] from the reference pattern and its rotated version. This can also serve as a tool for estimating the symmetry of images. That latter topic will be explored in Chapter 6.

The aforementioned favorable properties of the Zernike moments are valid as long as one uses a true analog image function. In practice, the Zernike moments have to be computed from sampled data. Thus let us define the following discrete version of A_{pq} over the pixel set $\{(x_i,y_j), 1 \leq i \leq n, 1 \leq j \leq n\}$ with the pixel width Δ

$$\widehat{A}_{pq} = \sum_{x_i^2+y_j^2 \leq 1}\sum h_{pq}(x_i,y_j)f(x_i,y_j), \tag{3.17}$$

where

$$h_{pq}(x_i,y_j) = \int_{x_i-\frac{\Delta}{2}}^{x_i+\frac{\Delta}{2}}\int_{y_j-\frac{\Delta}{2}}^{y_j+\frac{\Delta}{2}} V^*_{pq}(x,y)dxdy \tag{3.18}$$

represents the integration of $V^*_{pq}(x,y)$ over the (i,j) pixel. Since the image function is defined over the unit disk, the summation in (3.17) takes into account only those pixels whose centers fall completely inside the circle. It is important to observe that even if $f(x,y)=1$ for all $(x,y) \in D$, then

$$\widehat{A}_{pq} \neq A_{pq}. \tag{3.19}$$

This is in sharp contrast with the digital approximation for the geometric and Legendre moments for which there is equality in (3.19), see [144] and Chapter 2 in this monograph.

Hence, there is an inherent error in computing $\widehat{A}_{pq}$ related to the circular nature of the support of $\{V_{pq}(x,y)\}$. In what follows we refer to such an error as the geometric error. For a general image function, not necessarily constant, the error between $\widehat{A}_{pq}$ and A_{pq} can be decomposed as follows

$$E_{pq} = \widehat{A}_{pq} - A_{pq} = E^g_{pq} + E^n_{pq}, \tag{3.20}$$

where E^g_{pq} is the geometric error and E^n_{pq} is the numerical error related to the need of numerical integration in (3.18). The geometric and numerical error components of the overall error (3.20) are examined in the following section and further studied in Chapter 4 in the context of the image reconstruction problem.

3.4. Error Analysis

In this section, we examine the accuracy of the estimate $\widehat{A}_{pq}$. Two inherent sources of error in $\widehat{A}_{pq}$ are found, i.e., the geometric and numerical errors.

Geometric Error

In order to reveal the fundamental nature of the geometric error, let us consider the following expression

$$G(\Delta) = \Delta^2 \sum\sum_{x_i^2+y_j^2 \in D} 1 - \pi. \tag{3.21}$$

Let us observe that this is equal to $\widehat{A}_{00} - A_{00}$ provided that $f(x,y) = 1$ for all $(x,y) \in D$. In fact let us observe that $\iint_D V^*_{pq}(x,y)dxdy = \pi$, $V^*_{00}(x,y) = 1$ and that in our case the unit disk is located in the square $[-1,1]^2$ which is composed of n^2 pixels, i.e., $\Delta = 2/n$.

The term $G(\Delta)$ is not equal to zero due to the fact that if the center of a pixel falls inside the border of the unit disk $\{(x,y) : x^2 + y^2 \leq 1\}$, this pixel is used in the computation of the Zernike moments, otherwise, the pixel is discarded. Therefore, the area used for the moment computing is not equal to the area of the unit disk. Fig. 3.4 shows the union of the pixels whose centers fall inside the unit circle. Note that some pixels are not entirely inside the circle; on the other hand, some parts of the circle are not covered by the pixels.

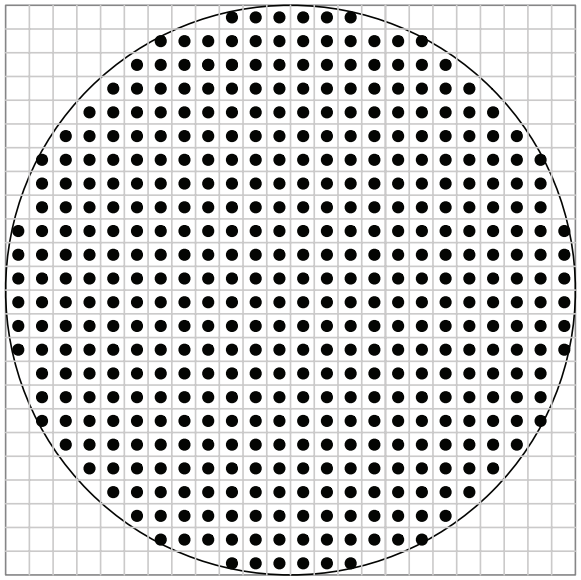


Figure 3.4: Lattice-point approximation of a circular region

The quantity $K(\Delta) = \sum_{x_i^2+y_j^2 \in D} 1$ in (3.21) denotes the number of the points $\{(x_i, y_j) : 1 \leq i, j \leq n\}$ inside the unit circle. Hence $G(\Delta)$ can be rewritten as follows

$$G(\Delta) = \Delta^2 K(\Delta) - \pi. \tag{3.22}$$

Formula (3.22) fully describes the nature of the geometric error and it is crucial to know the size of $G(\Delta)$, i.e., how fast $G(\Delta)$ tends to zero as $\Delta \to 0$. Since the term $\Delta^2 K(\Delta)$ can be approximated by an integral $\iint_D dxdy = \pi$, therefore the quantity $G(\Delta)$ in (3.22) plays the role of a remainder term of this approximation. The evaluation of the magnitude of $G(\Delta)$ as $\Delta \to 0$ is not a trivial problem. Nevertheless it turns out that quantity $G(\Delta)$ has been extensively examined in the analytic number theory with the relation to the so-called lattice points of a circle problem due originally to Gauss [104, 110, 119]. Gauss' problem concerning the number of points inside a circle is to determine the correct order of magnitude of $G(\Delta)$ as $\Delta \to 0$. First of all, it is known after Gauss that $G(\Delta) = O(\Delta)$. This result was improved by Sierpiński in 1906 to the form of $G(\Delta) = O(\Delta^{4/3})$. Generally $G(\Delta) = O(\Delta^{2(1-\theta)}), 1/4 < \theta \leq 1/2$ with the following significant steps in the history of finding the smallest possible θ:

- Gauss (1834), $\theta = 1/2 = 0.5$
- Sierpiński (1906), $\theta = 1/3 = 0.3333...$
- Walfisz (1927), $\theta = 163/494 = 0.329959514...$
- Titchmarsh (1935), $\theta = 15/46 = 0.326086957...$
- Hua (1942), $\theta = 13/40 = 0.325$
- Iwaniec and Mozzochi (1988), [110], $\theta = 7/22 = 0.318181818...$
- Huxley (2002), [104, 106], $\theta = 131/416 = 0.3149038...$

Hardy's conjecture says that the value of θ can be arbitrary close to $\theta = 1/4 = 0.25$. It is also known that $\theta = 1/4$ is impossible. This still remains an open problem in analytic number theory. See Chapter 4 for further discussion and [104, 106, 110, 119].

Hence we have the following result for the size of $G(\Delta)$

$$G(\Delta) = O(\Delta^{\gamma}), \quad 1 \leq \gamma < 3/2, \tag{3.23}$$

with $\gamma = 285/208$ due to the latest result of Huxley [106].

The relationship of $G(\Delta)$ to the geometric error in the decomposition proposed in (3.20) is described by the following result.

Theorem 3.2. *Let $f \in L_2(D)$ be a bounded function on D. Then for any admissible pair (p, q) we have*

$$|E_{pq}^{g}| \leq \overline{f} \min\left(\left\{ \frac{\pi}{p+1} \right\}^{\frac{1}{2}} |G(\Delta)|^{1/2}, |G(\Delta)| \right), \tag{3.24}$$

where $G(\Delta)$ is defined in (3.22) and $\overline{f} = \max_{x,y} f(x, y)$.

The proof of Theorem 3.2 can be found in the Appendix where we also give the precise formula for E_{pq}^{g}. Let us observe that the bound is controlled by the geometric factor $G(\Delta)$ and the degree p of the Zernike moment. Hence we have two terms in the bound for the geometric error. The first one is of order $\sqrt{G(\Delta)/p}$, whereas the second one is just $G(\Delta)$. The first term dominates if only $p > c\Delta^{-\gamma/4}$, c being some constant, assuming that $G(\Delta) = O(\Delta^{\gamma})$.

The rate of decaying of E_{pq}^{g} is limited by the best possible result for the lattice approximation of the unit circle. By virtue of Hardy's conjecture, the rate for E_{pq}^{g} can never exceed $O(\Delta^{3/4-\delta} \min(p^{-1/2}, \Delta^{3/4-\delta}))$ for an arbitrary small $\delta > 0$. With the best possible result known so far due to Huxley [106], the geometric error is of order

$$|E_{pq}^{g}| = O(\Delta^{285/416} \min(p^{-1/2}, \Delta^{285/416})). \tag{3.25}$$

Once again the bound is $O(p^{-1/2}\Delta^{285/416})$ if $p > n^{285/832} = n^{0.3425....}$, otherwise we have the bound $O(\Delta^{285/208})$.

Numerical Error

The numerical error is caused by the need of calculating accurately two dimensional integrals appearing in the definition of A_{pq}. The following result describes the size of the numerical error in the decomposition proposed in (3.20) for a wide class of image functions.

Theorem 3.3. *Let $f \in L_2(D)$ be a function of bounded variation on D. Then for any admissible pair (p, q) we have*

$$|E_{pq}^n| \le \left\{ \frac{\overline{f}\pi}{p+1} \operatorname{TV}(f) \right\}^{\frac{1}{2}} \Delta, \tag{3.26}$$

where $\operatorname{TV}(f)$ is the total variation of f, $\overline{f} = \max_{x,y} f(x, y)$.

The proof of Theorem 3.3 can be found in the Appendix where we also give the precise formula for E_{pq}^n. It is worth noting that the bound for the numerical error can be further improved if stronger than the bounded variation assumption on the image function $f(x, y)$ is imposed, e.g., that $f(x, y)$ has a number of derivatives. Then the rate of decreasing of E_{pq}^n can be faster than $O(\Delta)$. Nevertheless from practical point of view, we need to evaluate the factor $h_{pq}(x_i, y_j)$ defined in (3.18). For small values of q and p, this can be done by a direct integration in (3.18). For moderate and large values of p and q, some numerical integration techniques are needed.

The simplest strategy is the one dimensional integration rule to approximate $h_{pq}(x_i, y_j)$ given by

$$h_{pq}(x_i, y_j) \simeq \Delta^2 V_{pq}^*(x_i, y_j) \tag{3.27}$$

with the approximation error of order $O(\Delta^4)$.

Multidimensional integration rules of order L are of the following form

$$h_{pq}(x_i, y_j) \simeq \Delta^2 \sum_{l=1}^{L} w_l V_{pq}^*(u_l, v_l), \tag{3.28}$$

where $\{(u_l, v_l),\ 1 \le l \le L\}$ is a set of design points belonging to the (i, j) pixel centered at (x_i, y_j) and $\{w_l,\ 1 \le l \le L\}$ is a set of weights [40]. With an appropriate selection of the design points and weights, one can reduce the approximation error in (3.28) to $O\left(\Delta^{2(L+1)}\right)$ and therefore to achieve a subpixel accuracy.

For the simplicity of notation, let us denote by $(\pm 1, \pm 1)$ the corner points of the individual pixel. Then the 5-dimensional cubature formula for calculating $\int_{-1}^{1}\int_{-1}^{1} V(x,y)dxdy$ for some function $V(x,y)$ is given by

$$C_5(V) = \frac{1}{3}\{8V(0,0) + V(0,1) + V(1,0) + V(0,-1) + V(-1,0)\}. \quad (3.29)$$

This is illustrated in Fig. 3.5a.

(a) **(b)**

Figure 3.5: 5-dimensional cubature formulas I and II

Yet another type of the 5-dimensional cubature formula, which is shown in Fig. 3.5b, is defined as follows:

$$C_5(V) = \frac{4}{3}\{-V(0,0) + V(0,0.5) + V(0.5,0) + V(0,-0.5) + V(-0.5,0)\}. \quad (3.30)$$

The number of nodes in each pixel can be increased further to achieve even higher accuracy. An example is to use the 13-dimensional cubature formula, whose nodes are shown in Fig. 3.6.

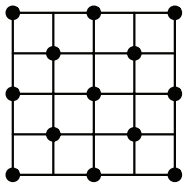


Figure 3.6: 13-dimensional cubature formula

Two different ways of calculating weights yield the following formulas referred to as the 13-D(I) and 13-D(II) integration methods, respectively

$$\begin{aligned} C_{13}(V) \;=\; & \frac{1}{45}\{120V(0,0) + 8[V(1,1) + V(1,-1) + V(-1,-1) + V(-1,1)] \\ & + 16[V(0.5,0.5) + V(0.5,-0.5) + V(-0.5,-0.5) + V(-0.5,0.5)] \\ & - 9[V(0,1) + V(1,0) + V(0,-1) + V(-1,0)]\}, \end{aligned} \quad (3.31)$$

and

$$\begin{aligned} C_{13}(V) &= \frac{1}{45}\{104V(0,0) + 4[V(1,1) + V(1,-1) + V(-1,-1) + V(-1,1)] \\ &\quad + 16[V(0.5,0.5) + V(0.5,-0.5) + V(-0.5,-0.5) + V(-0.5,0.5)] \\ &\quad - [V(0,1) + V(1,0) + V(0,-1) + V(-1,0)]\}. \end{aligned} \tag{3.32}$$

See [31] for further details on the above two-dimensional numerical integration techniques.

Generally the higher dimensional integration rules are more accurate for calculating two dimensional integrals. It should, however, be noted that this is not the case in our situation since we are dealing with circular area. In fact for multi-dimensional cubature formulas, some $V^*_{pq}(u_l, v_l)$ used in the computation of $h_{pq}(x_i, y_j)$ do not satisfy the restriction of $u_l^2 + v_l^2 \leq 1$. For example, there are 40, 16, and 140 design points falling outside the unit disk for 5-dimensional formula I, 5-dimensional formula II, and 13-dimensional formulas, respectively. Note that there is no such difficulty with the 1-dimensional rule in (3.27). Nevertheless, a certain form of truncation applied to the multi-dimensional integration formulas can lead to their substantial improvement. Indeed, let us redefine the Zernike moment in (3.17) in the following form

$$\widehat{A}_{pq}(\gamma) = \sum\sum_{x_i^2 + y_j^2 \leq 1-\eta} h_{pq}(x_i, y_j) f(x_i, y_j), \tag{3.33}$$

where η is a small adjustable parameter. Note that $\widehat{A}_{pq}(0) = \widehat{A}_{pq}$. The choice $\eta = \dfrac{\Delta}{\sqrt{2}}$ implies that only pixels completely falling into the unit circle are taken into account in (3.33). This, however, increases the geometric error, i.e., leads to poorer digital approximation of the circle. A less conservative choice takes the following form

$$\eta = \frac{\Delta}{2} + \varepsilon, \tag{3.34}$$

where ε is a small number, e.g., $\varepsilon = 0.0001$. With this choice, the number of the design points falling outside the unit disk will be reduced to 16, 0, 68 for the 5-dimensional formula I, 5-dimensional formula II, and 13-dimensional formulas, respectively. We will show the efficiency of the above modification of the Zernike moments in Section 5. Moreover, in Chapter 5 we are giving a high precision method for computing A_{pq}.

3.4.1. Image Reconstruction from Zernike Moments

Combining the results of Theorems 3.2 and 3.3 we can readily obtain the following bound for $|\widehat{A}_{pq} - A_{pq}|$.

Theorem 3.4. *Under the assumptions of Theorems 3.2 and 3.3 we have*

$$|\widehat{A}_{pq} - A_{pq}| \leq \left\{ \frac{\overline{f}\pi}{p+1} \operatorname{TV}(f) \right\}^{1/2} \Delta + \overline{f} |G(\Delta)|^{1/2} \min \left(\left\{ \frac{\pi}{p+1} \right\}^{\frac{1}{2}}, |G(\Delta)|^{1/2} \right). \quad (3.35)$$

Hence with the help of the best known result from the analytic number theory, see (3.25), we obtain

$$|\widehat{A}_{pq} - A_{pq}| = O(p^{-1/2}\Delta) + O\left(\Delta^{285/416} \min(p^{-1/2}, \Delta^{285/416}) \right). \quad (3.36)$$

For small fixed values of p the above bound is of order $O(\Delta)$, i.e., it is controlled by the discretization error. Nevertheless, as we have already mentioned this error can be greatly reduced by applying higher order numerical integration methods yielding the discrertization error of order $O(\Delta^L)$, $L \geq 2$. Hence we can conclude that the geometric error actually dominates the magnitude of $|\widehat{A}_{pq} - A_{pq}|$, i.e., we have

$$|\widehat{A}_{pq} - A_{pq}| = O(\Delta^{285/208}) \quad (3.37)$$

for all finite p.

As has already been mentioned, in order to assess the performance of a set of image descriptors, one can look at their reconstruction power. Using (3.12) and replacing A_{pq} with $\widehat{A}_{pq}$, we can define the following reconstruction algorithm

$$\widehat{f}_T(x,y) = \sum_{p=0}^{T} \sum_{q=-p}^{p} \tau_p \widehat{A}_{pq} V_{pq}(x,y), \qquad p - |q| = \text{even}, \quad (3.38)$$

where T is the truncation parameter informing us how many moments are taken into account.

The energy of the error image $\widehat{f}_T(x,y) - f(x,y)$ can serve as a natural performance measure for $\widehat{f}_T(x,y)$, i.e.,

$$\operatorname{Error}(\widehat{f}_T) = \iint_D |\widehat{f}_T(x,y) - f(x,y)|^2 dxdy. \quad (3.39)$$

Owing to Parseval's formula, we obtain the following decomposition for $\operatorname{Error}(\widehat{f}_T)$

$$\operatorname{Error}(\widehat{f}_T) = \sum_{p=0}^{T} \sum_{q=-p}^{p} \tau_p |\widehat{A}_{pq} - A_{pq}|^2 + \sum_{p=T+1}^{\infty} \sum_{q=-p}^{p} \tau_p |A_{pq}|^2 = D_T + F_T, \quad (3.40)$$

where $p-|q|=$ even. Note that the term F_T is a result of using a finite number of moments in the reconstruction algorithm. It is also clear that D_T increases with T, whereas F_T decreases with T. This reveals that there is T^* which minimizes an apparent trade-off between the terms in (3.40). A precise analysis of the reconstruction error will be given in Chapter 4.

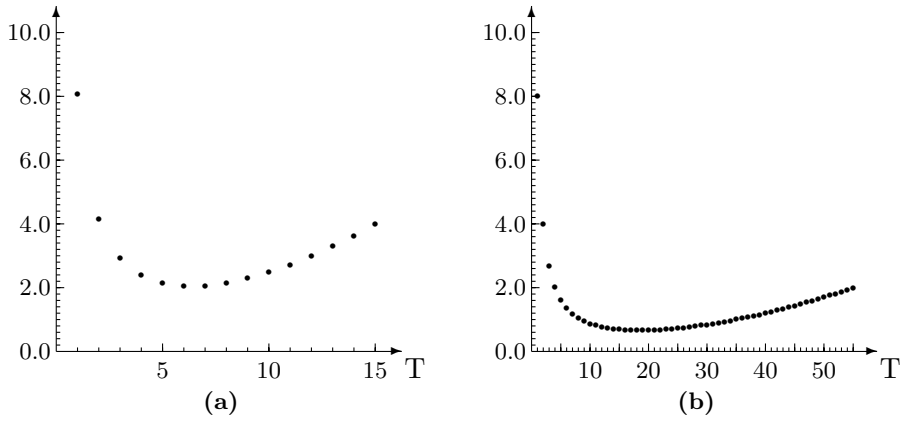


Figure 3.7: Error($\widehat{f}_T$) versus T for $a=1$, $b=0$, (a) $n=32$, (b) $n=256$

Instead of giving a general theory of the reconstruction problem we have evaluated the error formula in (3.40) for the following discontinuous image function of the radial form

$$f(x,y)=\begin{cases} a & x^2+y^2\le\eta^2, \\ b & \eta^2<x^2+y^2\le 1, \end{cases} \tag{3.41}$$

where $0<\eta<1$ and, without loss of generality, $a>b$.

In this case it can be calculated after tedious but otherwise straightforward algebra that the truncation error is bounded by

$$F_T \le \frac{8(a-b)^2}{T}. \tag{3.42}$$

Furthermore noting that $\overline{f}=a$ and $V(f)=a-b$, we have evaluated the first term in (3.40). Fig. 3.7 plots the bound versus T with $a=1, b=0$, and $n=32, n=256$. Note that T minimizing the bound for (3.40) is equal to 6 and 20 for $n=32$ and $n=256$, respectively. The optimal choice of T says how many moments should be taken into account in order to achieve the smallest possible reconstruction error.

3.5. Experimental Results

A set of five Chinese characters are employed as the test images. Each image is of the size of 24×24 pixels and the range of graylevels for each pixel is 32, where all characters have the gray level 11 and the background has the value 21. Fig. 3.8 illustrates these five Chinese characters projected onto the unit disk.

Figure 3.8: Five original Chinese characters used in image reconstruction via Zernike moments. From left to right are C_1, C_2, C_3, C_4, and C_5

3.5.1. Traditional Zernike Moment Method

Here we calculate A_{pq} employing the simplest 1-dimensional integration formula. The normalized mean square error

$$\overline{\text{Error}}(\widehat{f}_T) = \frac{\iint_D |f(x,y) - \widehat{f}(x,y)|^2 dxdy}{\iint_D [f(x,y)]^2 dxdy} \tag{3.43}$$

is adopted here as a measure of the accuracy of the reconstructed images.

Figure 3.9: Reconstructed patterns of the character C_1 based on the Zernike moments of order 14, 16, 18, 20, 22, 24, 26, 28, and 30

The Chinese character C_1 is used as the test image. Fig. 3.9 illustrates the reconstructed images of C_1. The first pattern is reconstructed from moments of order 14, then from left to right are the reconstructed images from moments of order 16, 18, 20, 22, 24, 26, 28, and 30, respectively.

3.5.2. Modified Zernike Moment Method

We have introduced a modified version $\widehat{A}_{pq}(\eta)$ of the Zernike moments in (3.33) of Section 3.2. In our experiment, the parameter η is chosen as $\eta = \Delta/2 + 0.0001$.

For the sake of comparison, the same Chinese character C_1 is employed as the test image, and the normalized mean square error defined in (3.43) is used.

Fig. 3.10 shows the $\overline{\text{Error}}(\widehat{f}_T)$ values for all five types of numerical integration formulas as a function of the moment order. Fig. 3.10 indicates that all five integration techniques perform better than the simple 1-dimensional integration formula in the traditional Zernike moment method. Among the applied integration techniques, the 5-dimensional formula II which is shown in Fig. 3.5 (b), is superior to the other techniques and therefore seems to be the best candidate for the image reconstruction for this specific situation.

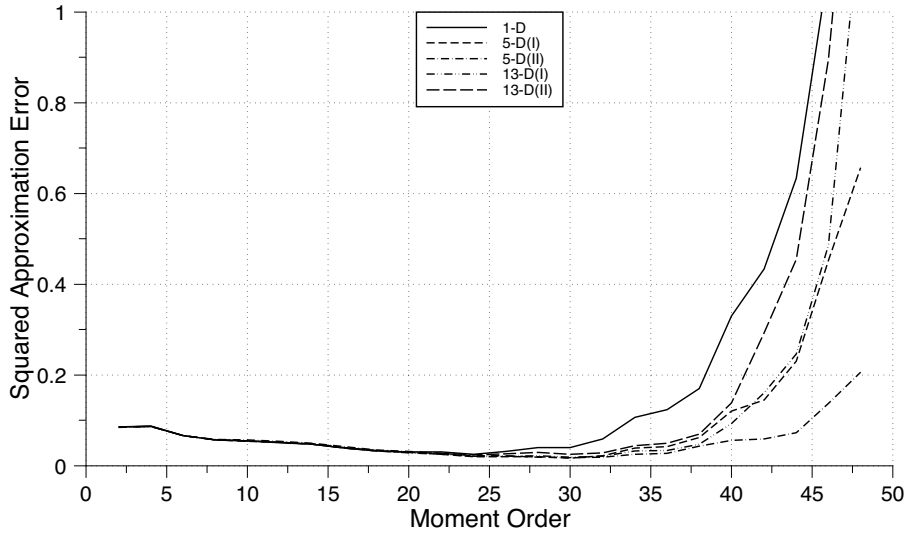


Figure 3.10: The normalized reconstruction error for the reconstructed character C_1 via the proposed Zernike moment technique with five numerical integration formulas

The reason why the 5-dimensional formula II provides better result is that with the new condition for the Zernike moments, all nodes used in this formula fall inside the unit disk. For the 5-dimensional formula I and the 13-dimensional formulas, we have 16 and 68 nodes falling outside the unit disk, respectively.

The reconstructed images of the character C_1 with five integration formulas are shown in Fig. 3.11. The first row shows the reconstructed patterns from the 1-dimensional integration formula, whereas the second, third, fourth, and fifth show 5-dimensional formulas I and II, 13-dimensional formulas I, and II, correspondingly. All images in the first column are reconstructed from moment of order 10, then from left to right we show results based on moments of order 15, 20, 25, 30, 35, and 40.

By using the 5-dimensional integration formula II, we reconstructed all five

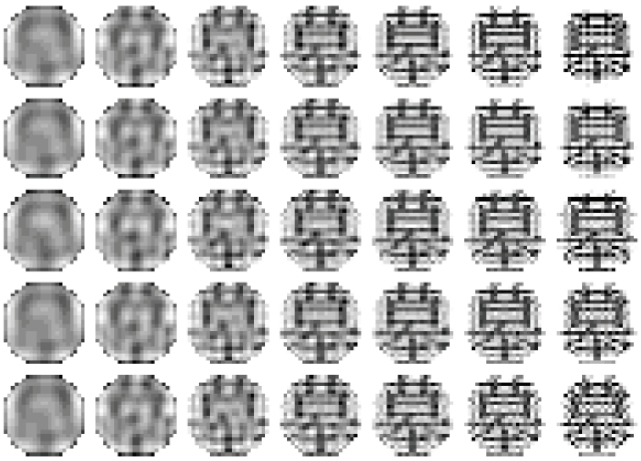


Figure 3.11: Reconstructed patterns of the character C_1 utilizing the modified method and five different integration rules. From left to right, the results based on moments of order 10, 15, 20, 25, 30, 35, and 40

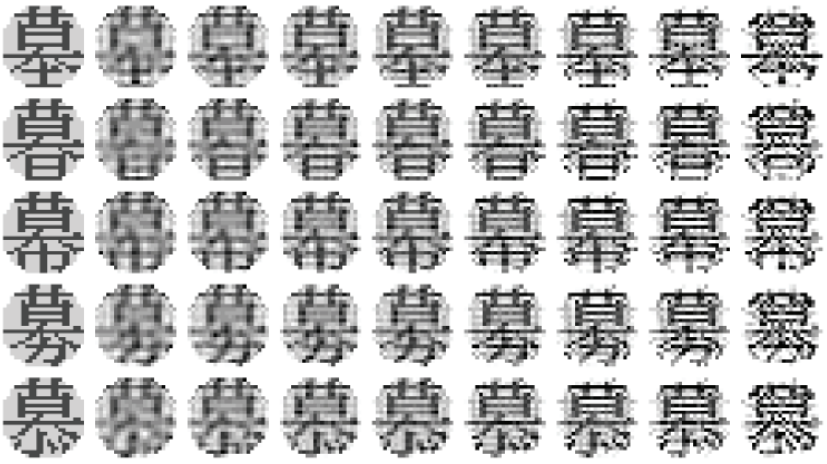


Figure 3.12: Utilizing the 5-dimensional integration formula II, the five Chinese characters and their reconstructed patterns via the modified Zernike moments of order 20, 24, 28, 32, 36, 40, 44, and 48

Chinese characters shown in Fig. 3.8 with the Zernike moments of orders from 20 to 48. Fig. 3.12 depicts the reconstructed images.

3.6. Conclusions

In this chapter, we have studied the accuracy problem for Zernike moments determined from discrete data observed on the $n \times n$ pixels image domain. It has been observed that two kinds of errors determine the accuracy of Zernike moments computing, i.e., the numerical error related to the usual need of calculating accurately two dimensional integrals and the geometric error being a distinctive feature of Zernike moments. We have shown that the geometric error plays a critical role for the accuracy of the Zernike moment computing. Employing the latest results from the analytic number theory on the lattice points of a circle, we have established that the accuracy of an estimate of the Zernike moment of order p is $O(\Delta^{285/416} \min(p^{-1/2}, \Delta^{285/416}))$. This result allows us to evaluate the integrated squared error of the reconstruction technique based on the maximum T estimated Zernike moments. This issue will be addressed in detail in Chapter 4.

The results established in this chapter can be further extended to other classes of radial polynomials of the invariant form. In fact an analogous theory can be carried out for the pseudo-Zernike polynomials introduced in Section 3.2. Furthermore, one can generalize the class of radial polynomials by introducing a tuning parameter which may yield a more flexible class of invariant radial moments. In fact the representation in (3.8) can be replaced by $\rho^{|q|} P_s^{(\alpha,|q|)}(2\rho^2 - 1)$, where $\alpha > -1$ is the tuning parameter. For $\alpha = 0$ we obtain the classical Zernike polynomials. The reconstruction properties of moments using such generalized orthogonal bases, however, are rather unknown and this is left for future studies.

In summary we have obtained in this chapter the following main results:

- Basic analytical and invariant properties of orthogonal radial polynomials have been established. In particular, the orthogonal radial Zernike polynomials have been examined.
- An improved technique for high quality calculation of Zernike moments has been proposed.
- The error analysis of the numerical estimate of the Zernike moment has been carried out.
- The fundamental connection between the geometric error and the analytic number theory of lattice points has been established.
- High order numerical integration methods for Zernike moments computing have been developed.
- Bounds for the precision of the proposed estimates have been established.

- Modified Zernike moments with the reduced geometric error have been proposed.
- Numerical studies have been carried out to illustrate our basic theory.

3.7. Appendix

Proof of Theorem 3.2 and Theorem 3.3

Let $D = \{(x, y) : x^2 + y^2 \leq 1\}$ be the unit disk. Let p_{ij} denote the (i, j) pixel at (x_i, y_j). Note that the single pixel has the area $4/n^2$, where n^2 is the total number of pixels.

Redefining the (p, q) Zernike polynomial as $V_{pq}(x, y)$ on D and zero otherwise, we can decompose A_{pq} as follows

$$A_{pq} = \sum_{x_i^2+y_j^2 \in D} \sum \iint_{p_{ij}} f(x, y) V_{pq}^*(x, y) dxdy + \iint_{O(\Delta)} f(x, y) V_{pq}^*(x, y) dxdy, \quad (3.44)$$

where $O(\Delta)$ is the region being the intersection of D with the union of those pixels whose centers and corners fall outside the circle.

Using (3.44), we can write the following decomposition for $\widehat{A}_{pq} - A_{pq}$

$$\begin{aligned} \widehat{A}_{pq} - A_{pq} &= \sum_{x_i^2+y_j^2 \in D} \sum \iint_{p_{ij}} \left(f(x_i, y_j) - f(x, y)\right) V_{pq}^*(x, y) dxdy \\ &\quad - \iint_{O(\Delta)} f(x, y) V_{pq}^*(x, y) dxdy. \end{aligned} \quad (3.45)$$

The first term of this decomposition defines the numerical error E_{pq}^n, while the second the geometric error E_{pq}^g, i.e.,

$$E_{pq}^n = \sum_{x_i^2+y_j^2 \in D} \sum \iint_{p_{ij}} \left(f(x_i, y_j) - f(x, y)\right) V_{pq}^*(x, y) dxdy, \quad (3.46)$$

$$E_{pq}^g = - \iint_{O(\Delta)} f(x, y) V_{pq}^*(x, y) dxdy. \quad (3.47)$$

Let us first consider the term E_{pq}^n. By virtue of Cauchy-Schwartz inequality for

integrals and sums, we have

$$|E_{pq}^n| \leq \sum_{x_i^2+y_j^2\in D}\sum \left\{ \iint_{p_{ij}\cap D} (f(x_i,y_j)-f(x,y))^2\,dxdy \iint_{p_{ij}} \left|V_{pq}^*(x,y)\right|^2 dxdy \right\}^{\frac{1}{2}}$$

$$\leq \left\{ \sum_{x_i^2+y_j^2\in D}\sum \iint_{p_{ij}\cap D} (f(x_i,y_j)-f(x,y))^2\,dxdy \iint_{D} \left|V_{pq}^*(x,y)\right|^2 dxdy \right\}^{\frac{1}{2}}.$$

Arguing as in the proof of Theorem 1 in [144], we can show that for f being a bounded variation function, we have

$$\sum_{x_i^2+y_j^2\in D}\sum \iint_{p_{ij}\cap D} (f(x_i,y_j)-f(x,y))^2\,dxdy \leq \frac{\overline{f}\,\mathrm{TV}(f)}{n^2},$$

where $\mathrm{TV}(f)$ is the total variation of f and $\overline{f}=\max_{x,y} f(x,y)$.

Furthermore, due to the normalization we have

$$\iint_D \left|V_{pq}^*(x,y)\right|^2 dxdy = \frac{\pi}{p+1}.$$

All these considerations lead to

$$|E_{pq}^n| \leq \left\{ \frac{\pi\overline{f}V(f)}{p+1}\Delta^2 \right\}^{\frac{1}{2}}. \tag{3.48}$$

This proves Theorem 3.3.

Let us now take the geometric error E_{pq}^g into account. By Cauchy-Schwartz inequality, we have

$$\left|E_{pq}^g\right| \leq \left\{ \iint_{O(\Delta)} f^2(x,y)dxdy \iint_D \left|V_{pq}^*\right|^2 dxdy \right\}^{\frac{1}{2}} \leq \overline{f}\left\{ \frac{\pi}{p+1}|O(\Delta)| \right\}^{\frac{1}{2}}, \tag{3.49}$$

where $|O(\Delta)|$ is the area of the set $O(\Delta)$. The discussion in Section 3.1 yields an important observation that

$$|O(\Delta)| \leq |G(\Delta)|,$$

where $G(\Delta)$ is the remainder term of the lattice approximation of D, see (3.22).

An alternative bound for E_{pq}^g results from (3.47) and the fact that $|V_{pq}(x,y)| \leq 1$, i.e., we have

$$E_{pq}^g \leq \overline{f}|O(\Delta)| \leq \overline{f}|G(\Delta)|.$$

This proves the bound given in Theorem 3.2. □

Chapter 4

Reconstruction Aspects of Orthogonal Radial Moments

In this chapter we consider the problem of image reconstruction from the orthogonal radial moments introduced in Chapter 3. An image function $f(x, y)$ defined on the unit disk is observed in the presence of noise at discrete points of a regular square grid. An estimate of $f(x, y)$ based on the Zernike orthogonal radial functions over the unit disk is proposed. This class of functions has a distinctive property of being invariant to rotation of axes about the origin of coordinates yielding therefore a rotationally invariant estimate. For radial functions the orthogonal set has a particularly simple form being related to the classical Legendre polynomials. We give the statistical accuracy analysis of the proposed estimate of $f(x, y)$ in the sense of the L_2 metric. It is found that there is an inherent limitation in the precision of the estimate due to the geometric nature of a circular domain. This is explained by relating the accuracy issue to the celebrated problem in the analytic number theory called the lattice points of a circle. The accuracy of the reconstruction method for images with certain symmetry properties, like radiality, is examined.

In summary, the main contributions of this chapter are the following:

- Basic approximation properties of the Zernike orthogonal radial polynomials are established.
- The behavior of Zernike moments for images with various symmetry and structural conditions is extensively examined. This includes radial images and images of the additive form.
- An image reconstruction formula from the Zernike orthogonal radial polynomials is introduced and its reconstruction power is thoroughly examined.

- The detailed error analysis of the reconstruction algorithm is carried out. This includes the noise diminishing property and the imprecision of the Zernike moment calculation due to the geometric and discretization errors.
- The fundamental bound for the truncation reconstruction error is established. The bound holds for all images which belong to a class of functions of bounded variation.
- The rates for the mean integrated squared error are derived. The rates for important special cases of radial and additive images are given.
- The problem of data-driven selection of the optimal number of moments is discussed.

4.1. Introduction

Let us consider the image model introduced in Chapter 3, i.e., let $f(x, y)$ be an image function of the class $L_2(D)$ defined over the unit disk $D = \{(x, y) : x^2 + y^2 \leq 1\}$. In this chapter we address the problem of recovering $f(x, y)$ from the orthogonal radial moments examined in Chapter 3 when only noisy and discrete data are available. Hence let

$$Z_{ij} = f(x_i, y_j) + \varepsilon_{ij}, \ (x_i, y_j) \in D, \ 1 \leq i, j, \leq n \tag{4.1}$$

be the observed data record, where $\{\varepsilon_{ij}\}$ is a zero mean, finite variance, spatially uncorrelated noise process. Let σ^2 denote the variance of ε_{ij}. We assume that the data are observed on a square grid of edge width Δ, i.e., $x_i - x_{i-1} = y_j - y_{j-1} = \Delta$.

As we have already noted we consider functions on the circular domain since it is a common situation in a wide range of applications. Nevertheless, no theory on the reconstruction accuracy of radial moments in such a setting has been develolped so far. In this chapter we propose an estimation technique for recovering $f \in L_2(D)$ utilizing the Zernike orthogonal basis on D. Although it is possible to construct a number of orthogonal systems on D [126], the Zernike orthogonal functions [13, 18, 33, 250] possess some unique properties. First of all they are invariant in form with respect to orthogonal transformations of the (x, y)-plane yielding a popular class of invariant features in object recognition [1, 101, 121, 124, 127, 144, 145, 173, 199, 219, 221, 228, 232]. Second the Zernike functions have properties similar to both classical trigonometric series and Jacobi polynomials and they define a minimal set of permissible functions [13]. Furthermore, Radon transforms of the Zernike functions are also orthogonal [33]. The latter plays a fundamental role in tomographic reconstruction of objects from their projections [33, 115, 150, 157]. It is also worth mentioning that the Zernike

orthogonal basis has recently found applications in ophthalmology for corneal imaging allowing eye-care providers to fit contact lenses and glasses [207].

In this chapter the accuracy analysis of the proposed estimate of $f(x,y)$ based on Zernike functions is carried out. Let $\widehat{f}_T(x,y)$ be an estimate of $f(x,y)$ derived from data record (4.1) which utilizes the first T Zernike functions. We give bounds for the mean integrated squared error (MISE)

$$\mathrm{MISE}(\widehat{f}_T) = E \iint_D \left| \widehat{f}_T(x,y) - f(x,y) \right|^2 dxdy \tag{4.2}$$

for a large class of functions belonging to $L_2(D)$. This includes bounded variation functions, radial functions, and functions of a specific structural form. This analysis reveals the dependence of the error on the truncation point T, image smoothness, noise characteristics, sampling rate Δ and the circular geometry of the support of $f(x,y)$. This is quantified by decomposing the $\mathrm{MISE}(\widehat{f}_T)$ into the variance, bias, discretization and geometric error components. The first three factors are standard in the non-parametric curve estimation theory and yield the well studied variance/bias/discretization trade-off [39, 210]. The geometric error is caused by the fact that our reconstruction problem is confined to a unit circle and some lattice squares along the circular domain may be either included or excluded. This kind of error is quantified by using the celebrated problem in the analytic number theory referred to as lattice points of a circle due originally to Gauss [87, 104–106, 110, 235]. We show that the geometric error, although not dominant in the asymptotic theory, has an important influence on the overall performance of our reconstruction method. This is a distinctive feature of the function reconstruction problem on a circular domain.

4.2. Orthogonal Bases on D

In this sub - section we give an essential information about the Zernike orthogonal functions which will be employed further. A detailed discussion of this subject can be found in Chapter 3.

First of all, an important property of the Zernike orthogonal functions is that the radial polynomial $R_{pq}(\rho)$ is closely related to the classical Jacobi polynomials [13, 33, 126, 217] according to the relation

$$R_{pq}(\rho) = \rho^{|q|} P_s^{(0,|q|)}(2\rho^2 - 1), \tag{4.3}$$

where $s = (p - |q|)/2$ and $P_n^{(\alpha,\beta)}(x)$ is the Jacobi polynomial of order n with the parameter (α, β).

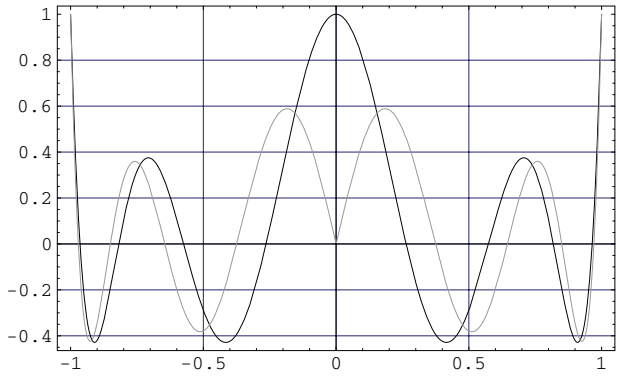


Figure 4.1: Radial section of $R_{pq}(\rho)$ for $(p,q) = (8,0)$ (bold line) and $(p,q) = (9,1)$ (light line)

The identity in (4.3) implies that if $p = |q|$ then

$$R_{pp}(\rho) = \rho^p, \tag{4.4}$$

while for $q = 0$ we have

$$R_{p0}(\rho) = P_{p/2}(2\rho^2 - 1), \tag{4.5}$$

where $P_s(x)$ is the s-th order classical Legendre polynomial defined on $[-1,1]$, [206, 217]. It is also known [205] that

$$|P_n(x)| < 1, -1 < x < 1. \tag{4.6}$$

For our further considerations it is useful to know the growth of the polynomial $R_{pq}(\rho)$. First let us recall the fact that $P_n^{(\alpha,\beta)}(x)$ attains its maximum in $[-1,1]$ at one of the end-points [44, p. 168]. Hence observing that $P_s^{(0,|q|)}(1) = 1$,

$$P_s^{(0,|q|)}(-1) = (-1)^s \binom{s+q}{s} \simeq s^q$$

and then using (4.3) and (4.6), we obtain the following useful bound for $R_{pq}(\rho)$

$$|R_{pq}(\rho)| \leq 1, \tag{4.7}$$

with $R_{pq}(1) = 1$.

Fig. 4.1 confirms the above bounds by showing the radial section of $R_{pq}(\rho)$ for $(p,q) = (8,0)$ and $(p,q) = (9,1)$.

4.3. Function Approximation by Zernike Functions

The completeness and orthogonality of $\{V_{pq}(x,y)\}$ allow us to represent any $f \in L_2(D)$ by the following series

$$\sum_{p=0}^{\infty} \sum_{q=-p}^{p} \tau_p A_{pq} V_{pq}(x,y), \tag{4.8}$$

where the summation is carried out for all (p,q) satisfying (3.12) and where, due to (3.4), $\tau_p = (p+1)/\pi$ is the normalizing constant. The Fourier coefficient A_{pq} of order p with repetition q is given by

$$A_{pq} = \iint_D f(x,y) V_{pq}^*(x,y) dx dy. \tag{4.9}$$

Recalling (3.3), we can obtain the following useful polar coordinate version of (4.9)

$$A_{pq} = \int_0^{2\pi} \int_0^1 f(\rho\cos\theta, \rho\sin\theta) R_{pq}(\rho) e^{-jq\theta} \rho d\rho d\theta. \tag{4.10}$$

Denoting $\widetilde{f}(\rho,\theta) = f(\rho\cos\theta, \rho\sin\theta)$ and observing that $\widetilde{f}(\rho,\theta)$ is 2π-periodic in θ we can rewrite (4.10) in the following equivalent form

$$A_{pq} = 2\pi \int_0^1 c_q(\rho) R_{pq}(\rho) \rho d\rho, \tag{4.11}$$

where

$$c_q(\rho) = \frac{1}{2\pi} \int_0^{2\pi} \widetilde{f}(\rho,\theta) e^{-jq\theta} d\theta \tag{4.12}$$

is the usual q-th Fourier coefficient of the function $\varphi(\theta) = \widetilde{f}(\rho,\theta)$.

The above identities allow us to draw inference about the coefficient A_{pq} under different symmetry conditions put on the image function $f(x,y)$. This issue was discussed in Section 3.3 and will be further examined in Chapter 6.

It is important for further studies to evaluate the form of A_{pq} for various classes of functions defined on D. Let us first consider an important class of radial functions, i.e., let $f(x,y)$ be a function of $\rho^2 = x^2 + y^2$ only; write it as $g(\rho^2)$. In this case it is easy to observe that $c_q(\rho)$ defined in (4.12) is given by

$$c_q(\rho) = g(\rho^2) \mathbf{1}(q=0),$$

where $\mathbf{1}(A)$ denotes the indicator function of a set A.

By this and (4.5) we have

$$A_{pq} = 2\pi \int_0^1 g(\rho^2) P_{p/2}(2\rho^2 - 1)\rho d\rho \mathbf{1}(q = 0). \tag{4.13}$$

More general case to the above is when $f(x, y)$ is elliptically contoured, i.e.,

$$f(x, y) = g\left((x, y)\Sigma^{-1}\begin{pmatrix} x \\ y \end{pmatrix}\right) \tag{4.14}$$

for some non-singular matrix Σ and a single variable function $g(t)$. In particular let

$$\Sigma = \begin{pmatrix} 1 & \tau \\ \tau & 1 \end{pmatrix}$$

where τ plays the role of the "correlation" coefficient between the variables x and y. Assuming $|\tau| < 1$, we observe that (4.14) becomes

$$f(x, y) = g\left(\frac{x^2 + y^2 - 2xy\tau}{1 - \tau^2}\right). \tag{4.15}$$

Note that for $\tau = 0$ the Zernike coefficient A_{pq} of function (4.14) is given by (4.13). For general τ, A_{pq} is defined in (4.11) with

$$c_q(\rho) = \frac{1}{2\pi} \int_0^{2\pi} g\left(\frac{\rho^2(1 - \tau \sin 2\theta)}{1 - \tau^2}\right) e^{-jq\theta} d\theta. \tag{4.16}$$

It is interesting to know the behavior of A_{pq} as a function of τ. Expanding (4.16) into Taylor series about $\tau = 0$ and after some tedious algebra we can obtain that

$$A_{pq} = a_p + \tau^2 b_p + O(\tau^4), \tag{4.17}$$

where

$$a_p = 2\pi \int_0^1 g(\rho^2) R_{p0}(\rho)\rho d\rho$$

and

$$b_p = \pi \int_0^1 \left[\frac{1}{2} g^{(2)}(\rho^2)\rho^2 + 2g^{(1)}(\rho^2)\right] R_{p0}(\rho)\rho^3 d\rho - \frac{\pi}{2} \int_0^1 g^{(2)}(\rho^2) R_{p4}(\rho)\rho^5 d\rho.$$

Here it has been assumed that $g(t)$ possesses two continuous derivatives. Clearly the term a_p is identical to (4.13) and as we have already explained, it represents

the Zernike coefficient for radial functions. It is interesting to observe that expansion (4.17) does not possess a linear term with respect to τ. Generally all terms with odd powers of τ are zero, provided that corresponding higher order derivatives of $g(t)$ exist.

Let us now assume that $g(t)$ is Lipschitz, i.e., that

$$|g(u) - g(v)| \leq M|u - v|^{\alpha}, \text{ for } M > 0 \text{ and } 0 < \alpha \leq 1. \tag{4.18}$$

Then it can be easily shown that

$$|A_{pq} - a_p| \leq 2\pi M \left(\frac{2|\tau|}{1-\tau^2}\right)^{\alpha} \int_0^1 |R_{pq}(\rho)|\rho^{2\alpha+1} d\rho. \tag{4.19}$$

This result gives the bound for the distance between the Zernike coefficients of functions defined in (4.15) and the corresponding coefficients of radial function $g(\rho^2)$.

Result (4.19) can be viewed as an example of more general concept of the radial approximation of a given class of functions. Hence let $\mathcal{R}(D) \subset L_2(D)$ be a class of radial functions in $L_2(D)$. Then it can be shown [247] that the orthogonal projection of $L_2(D)$ onto $\mathcal{R}(D)$ is of the form

$$h(\rho^2) = \frac{1}{2\pi} \int_0^{2\pi} f(\rho\cos\theta, \rho\sin\theta) d\theta. \tag{4.20}$$

Hence $h(\rho^2)$ represents the best radial approximation of $f \in L_2(D)$ and consequently there is a unique decomposition

$$f = h + r, \tag{4.21}$$

where $h \in \mathcal{R}(D)$ and r is orthogonal to h, i.e.,

$$\int_0^{2\pi} \int_0^1 h(\rho^2) r(\rho, \theta) \rho d\rho d\theta = 0.$$

Remark 4.1. *It is interesting to note that the orthogonal projection of the invariant function $V_{pq}(x, y)$ in (3.3) onto $\mathcal{R}(D)$ is given by $R_{p0}(\rho)$.*

It is of some interest to evaluate the size of the term r in (4.21). For the function class defined in (4.15), it can be easily shown that

$$\| f - h \| \leq M \sqrt{\frac{\pi}{2\alpha+1}} \left(\frac{2\tau}{1-\tau^2}\right)^{\alpha},$$

where the function $g(\rho^2)$ in (4.15) satisfies (4.18). Here $\|\cdot\|$ is the $L_2(D)$ norm.

Yet another important functional class [37] consists of functions which can be decomposed into separate angle and radial components of the following form

$$\widetilde{f}(\rho,\theta)=\sum_{s=1}^{L}\varphi_s(\rho^2)\psi_s(\theta), \tag{4.22}$$

for some measurable functions $\varphi_s(\rho^2)$, $\psi_s(\theta)$, $s=1,...,L$.

A simple algebra shows that the Zernike coefficients for this class of functions have the following form

$$A_{pq}=2\pi\sum_{s=1}^{L}c_{s,q}B_{pq}(s), \tag{4.23}$$

where $c_{s,q}$ is the q-th Fourier coefficient (see (4.12)) of the angle component $\psi_s(\theta)$ and

$$B_{pq}(s)=\int_0^1\varphi_s(\rho^2)R_{pq}(\rho)\rho d\rho.$$

It can be easily shown that for the orthogonal projection $h(\rho^2)$ of class (4.22) onto $\mathcal{R}(D)$ we have the following bound

$$\|f-h\|\leq M\frac{(2\pi)^{\alpha+1/2}}{(\alpha+1)}\left\{\int_0^1\left(\sum_{s=1}^{L}\varphi_s(\rho^2)\right)^2\rho d\rho\right\}^{1/2},$$

where it is assumed that all $\psi_s(\theta), s=1,...,L$ satisfy (4.18). On the other hand, a bound in terms of the global L_2 variation of $\psi_s(\theta), s=1,...,L$ is of the form

$$\|f-h\|\leq 4\sqrt{2/3}\left\{\int_0^1\left(\sum_{s=1}^{L}\|\psi_s^{(1)}\|\,|\varphi_s(\rho^2)|\right)^2\rho d\rho\right\}^{1/2}. \tag{4.24}$$

For the proof of this result see [180] .

Remark 4.2. *The class of functions defined in (4.22) is important since it is known that it defines a dense subset in $L_2([0,2\pi]\times[0,1])$. In fact the set of all finite linear combinations*

$$\sum_{s=1}^{L}\varphi_s(\rho^2)\psi_s(\theta),\varphi_s\in L_2([0,1]),\psi_s\in L_2([0,2\pi])$$

is dense in $L_2([0,2\pi]\times[0,1])$.

A special interesting case of class (4.22) is the additive model

$$\widetilde{f}(\rho,\theta)=\varphi(\rho^2)+\psi(\theta) \tag{4.25}$$

for some measurable functions $\varphi(\rho^2)$ and $\psi(\theta)$. See [97] for an extensive overview of the theory and applications of such models in the statistical inference. It should be noted that both (4.21) and (4.25) describe additive models in the polar coordinate system. This should be contrasted with a counterpart of (4.25), $f(x,y)=f_1(x)+f_2(y)$, in the (x,y) coordinates.

The Zernike coefficient for $\widetilde{f}(\rho,\theta)$ in (4.25) takes the following form

$$A_{pq}=\mathbf{1}(q=0)2\pi\int_0^1\varphi(\rho^2)R_{p0}(\rho)\rho d\rho+\int_0^{2\pi}\psi(\theta)e^{-jq\theta}d\theta\int_0^1 R_{pq}(\rho)\rho d\rho. \tag{4.26}$$

It is also clear that the orthogonal projection of class (4.25) onto $\mathcal{R}(D)$ is given by

$$h(\rho^2)=\varphi(\rho^2)+\frac{1}{2\pi}\int_0^{2\pi}\psi(\theta)d\theta.$$

The distance between f in (4.25) and $h(\rho^2)$ is bounded by

$$\|\,f-h\,\|\leq\frac{4}{\sqrt{3}}\,\|\,\psi^{(1)}\,\|\,.$$

The proof of this bound is similar to (4.24). The interpretation of the above inequality and that in (4.24) is that if oscillations in the angular components $\{\psi_s\}$ are small then f is closer to h.

Let us also mention the class of ridge functions with L directions $\delta_0,...,\delta_{L-1}$

$$f(x,y)=\sum_{l=0}^{L-1}g_l(x\cos(\delta_l)+y\sin(\delta_l)), \tag{4.27}$$

where $\{g_l(t),l=0,...,L-1\}$ are some single variable functions defined on $[-1,1]$.

This class of functions plays an important role in projection-based approximation of multidimensional functions [37,150]. In particular it has been shown [150] that any polynomial of degree $L-1$ in x and y can be represented by (4.27) provided that $\{g_l(t),l=0,...,L-1\}$ are single variable polynomials of degree $L-1$ and $\delta_0,...,\delta_{L-1}$ are distinct directions. Furthermore the ridge functions can be used as a powerful tool for approximating radial functions and functions with smooth angular behavior [37].

It is straightforward to show that the Zernike coefficient for (4.27) takes the following form

$$A_{pq} = \sum_{l=0}^{L-1} e^{-jq\delta_l} \int_0^1 \left\{ \int_0^{2\pi} g_l(\rho\cos\theta)\cos(q\theta)d\theta \right\} \rho R_{pq}(\rho)d\rho.$$

This formula can be further simplified assuming a specific form of ridge functions $\{g_l(t), l = 0, ..., L-1\}$. In particular for $\{g_l(t), l = 0, ..., L-1\}$ being polynomials of degree not greater than $L-1$, it can be easily shown that $A_{pq} = 0$ for $q \geq L$.

Due to (4.20) we also note that the orthogonal projection of the class of ridge functions onto $\mathcal{R}(D)$ is of the form

$$h(\rho^2) = \frac{1}{2\pi} \sum_{l=0}^{L-1} \int_0^{2\pi} g_l(\rho\cos(\theta - \delta_l))d\theta.$$

For $g_l(t)$ satisfying (4.18), the L_2 distance between $h(\rho^2)$ and $f(x, y)$ in (4.27) is bounded by

$$\| f - h \| \leq \frac{LM(2\pi)^{\alpha+1/2}}{\sqrt{2}(\alpha+1)^{3/2}}.$$

On the other hand, a bound on $\| f - h \|$ in terms of the global variability of $\{g_l(t)\}$ is given by

$$\| f - h \| \leq 4\sqrt{2/3} \left\{ \int_0^1 \left| \sum_{l=0}^{L-1} \left(\int_0^{2\pi} \left| g_l^{(1)}(\rho\cos(\gamma - \alpha_l)) \right|^2 d\gamma \right)^{1/2} \right|^2 \rho^3 d\rho \right\}^{1/2}.$$

The proof of this fact can be obtained in the analogous way as the proof of (4.24).

To illustrate the richness of the above classes of image functions we give in Fig. 4.2 some examples. Hence we draw the density plots of the radial image function $f_1(x, y) = \sin(10\rho) + \cos(10\rho)$, the additive image function in the polar coordinates $f_2(x, y) = \sin(10\rho) + \cos(10\theta)$, and the additive image function in the Cartesian coordinates $f_3(x, y) = \sin(10x) + \cos(10y)$. Our developments in this chapter reveal that the Zernike based reconstruction method can easily recover functions like $f_1(x, y)$ and $f_2(x, y)$. The function $f_3(x, y)$ on the other hand would be more suitable for the Lagrange moments based reconstruction algorithms.

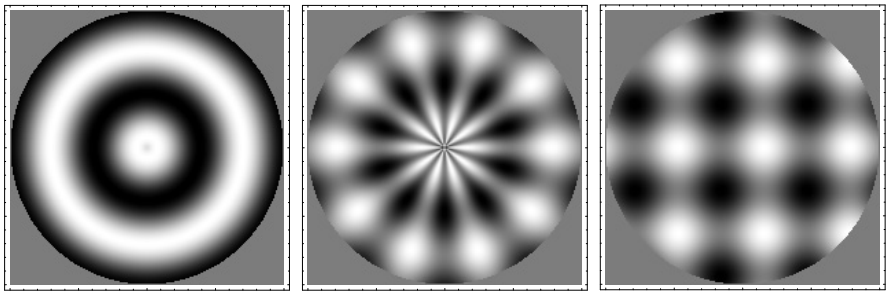


Figure 4.2: Examples of image functions: radial image function, additive image function in the (ρ, θ)-coordinates, additive image function in the (x, y)-coordinates

4.4. The Reconstruction Algorithm

Let us now consider the problem of estimating $f(x, y)$, where $f \in L_2(D)$, given noisy data generated from observation model (4.1). Let us assume that lattice points $\{(x_i, y_j), 1 \leq i, j \leq n\}$ from D are located at the center of the lattice squares. Fig. 3.4 in Section 3.4 shows the geometry of lattice points.

The first task is to estimate the Zernike coefficients A_{pq} defined in (4.9). Using a piecewise approximation of $f(x, y)$ over the lattice squares we can define the following estimate of A_{pq}

$$\widehat{A}_{pq} = \sum_{(x_i,y_j)\in D} \sum w_{pq}(x_i, y_j) Z_{ij}, \tag{4.28}$$

where

$$w_{pq}(x_i, y_j) = \int_{x_i-\frac{\Delta}{2}}^{x_i+\frac{\Delta}{2}} \int_{y_j-\frac{\Delta}{2}}^{y_j+\frac{\Delta}{2}} V_{pq}^*(x, y) dx dy.$$

It is also assumed that $V_{pq}(x, y)$ is defined on D by (3.3), (3.7), and is zero otherwise. Let us observe that

$$E\widehat{A}_{pq} = \sum_{(x_i,y_j)\in D} \sum w_{pq}(x_i, y_j) f(x_i, y_j)$$

is just a numerical approximation of the integral representing A_{pq} over the lattice.

It is worth noting that if $f(x, y) = 1$ for all $(x, y) \in D$ then we have $E\widehat{A}_{pq} \neq A_{pq}$. Such an unusual phenomenon does not hold for functions defined over the rectangular domain. In our case we are dealing with the circular domain and the lattice squares cannot fill the circle perfectly. This problem will be examined in detail in Section 4.5.

Representation (4.8) and formula (4.28) yield the following estimate of $f(x,y)$, $(x,y) \in D$

$$\widehat{f}_T(x,y) = \sum_{p=0}^{T} \sum_{q=-p}^{p} \tau_p \widehat{A}_{pq} V_{pq}(x,y), \tag{4.29}$$

for $p - |q|$ being even, $\tau_p = (p+1)/\pi$.

Estimate (4.29) is a standard orthogonal series estimate defined by a truncation parameter T. Let us observe that there are $(T+1)(T+2)/2$ terms in (4.29).

Remark 4.3. *Let us note that due to the invariant form of $V_{pq}(x,y)$ our estimate is rotationally invariant. Hence when any rotation $x' = x\cos(\phi) + y\sin(\phi)$, $y' = -x\sin(\phi) + y\cos(\phi)$ is applied, the estimate $\widehat{f}_T(x',y')$ is transformed into an estimate of the same form, i.e.,*

$$\widehat{f}_T(x',y') = \sum_{p=0}^{T} \sum_{q=-p}^{p} \tau_p \widehat{A}_{pq} e^{-jq\phi} V_{pq}(x,y).$$

To the best of author's knowledge no classical curve estimation method possesses this useful property.

For the class of radial functions $f(x,y) = g(\rho^2)$ considered in Section 4.3, our estimate can have a reduced complexity due to formula (4.13). In fact we can define the following estimate of $g(\rho^2)$

$$\widehat{g}_T(\rho^2) = \sum_{l=0}^{T} \tau_{2l} \widehat{A}_{2l} R_{2l}(\rho), \tag{4.30}$$

where

$$\widehat{A}_{2l} = 2\pi \sum_{i=1}^{n} \widetilde{Z}_i \int_{\rho_i - \frac{\Delta}{\sqrt{2}}}^{\rho_i + \frac{\Delta}{\sqrt{2}}} R_{2l}(\rho)\rho d\rho \tag{4.31}$$

is the estimate of A_{pq} in (4.13) with $p = 2l$. Here we denote $R_{2l}(\rho) = P_l(2\rho^2 - 1)$, where $P_l(t)$ is the l-th Legendre polynomial.

Furthermore $\{\widetilde{Z}_i\}$ are observations at the lattice points taken along the diagonal direction $\rho_i = \sqrt{x_i^2 + y_i^2} \leq 1$. Hence the actual number of points taken into account is of order n. It is also clear that one can use other directions like $\rho_i = |x_i|$ or $\rho_i = |y_i|$.

Remark 4.4. *In order to utilize the radial invariance of $f(x,y) = g(\rho^2)$ we can group data points according to their distance from the origin. Such groups define the ring regions dividing the disk. The data from a given group could be summarized by, e.g., their average and this would give an alternative definition of $\{\widetilde{Z}_i\}$ used in obtaining an estimate of A_{2l}. This approach uses all the data points and leads to the reduced variance of the corresponding estimate $\widehat{A}_{2l}$.*

The performance of $\widehat{f}_T(x,y)$ is measured by $\mathrm{MISE}(\widehat{f}_T)$ defined in (4.2). By virtue of (3.4) and Parseval's formula we obtain the following decomposition of $\mathrm{MISE}(\widehat{f}_T)$

$$\mathrm{MISE}(\widehat{f}_T) = \mathrm{VAR}(\widehat{f}_T) + \mathrm{BIAS}^2(\widehat{f}_T),$$

where

$$\begin{aligned} \mathrm{VAR}(\widehat{f}_T) &= \iint_D E\left|\widehat{f}_T(x,y) - E\widehat{f}_T(x,y)\right|^2 dxdy \\ &= \sum_{p=0}^{T}\sum_{q=-p}^{p} \tau_p E|\widehat{A}_{pq} - E\widehat{A}_{pq}|^2 \end{aligned} \tag{4.32}$$

is the integrated variance, whereas

$$\begin{aligned} \mathrm{BIAS}^2(\widehat{f}_T) &= \iint_D \left|E\widehat{f}_T(x,y) - f(x,y)\right|^2 dxdy \\ &= \sum_{p=0}^{T}\sum_{q=-p}^{p} \tau_p \left|E\widehat{A}_{pq} - A_{pq}\right|^2 + \sum_{p=T+1}^{\infty}\sum_{q=-p}^{p} \tau_p |A_{pq}|^2 \\ &= \mathrm{DG}(\widehat{f}_T) + \mathrm{TR}(\widehat{f}_T) \end{aligned} \tag{4.33}$$

is the integrated squared bias.

The term $\mathrm{TR}(\widehat{f}_T)$ is a bias caused by the truncation at T. The term $\mathrm{DG}(\widehat{f}_T)$ measures the discretization and geometric errors in estimating the coefficient A_{pq}. We show in Section 4.5 that $\mathrm{DG}(\widehat{f}_T)$ can be further decomposed into the discretization error $D(\widehat{f}_T)$ that reflects a usual need of numerical approximation of the integral and the geometric error $G(\widehat{f}_T)$ being a distinctive feature of our estimation problem. The geometric error measures the accuracy of the lattice approximation of the circular domain. It turns out that this approximation is described by a celebrated problem in the analytic number theory on the number of lattice points inside a circle [87, 105]. The contribution of each term in (4.32) and (4.33) to the overall error $\mathrm{MISE}(\widehat{f}_T)$ is examined in Section 4.5.

Concerning the estimate $\widehat{g}_T(\rho^2)$ in (4.30) and (4.31), let us first observe that

$$\mathrm{MISE}(\widehat{g}_T) = 2\pi \int_0^1 \left|\widehat{g}_T(\rho^2) - g(\rho^2)\right|^2 \rho d\rho.$$

This is decomposed as follows

$$\begin{aligned} \mathrm{MISE}(\widehat{g}_T) &= \sum_{l=0}^{T} \tau_{2l} E|\widehat{A}_{2l} - E\widehat{A}_{2l}|^2 + \sum_{l=0}^{T} \tau_{2l}|E\widehat{A}_{2l} - A_{2l}|^2 + \sum_{l=T+1}^{\infty} \tau_{2l}|A_{2l}|^2 \\ &= \mathrm{VAR}(\widehat{g}_T) + D(\widehat{g}_T) + \mathrm{TR}(\widehat{g}_T), \end{aligned} \tag{4.34}$$

where A_{2l} is A_{pq} in (4.13) with $p = 2l$.

It should be noted that in the case of radial functions the geometric error does not appear.

The estimation problem studied in this book falls into the category of non-parametric curve estimation since we do not assume any parametric knowledge of a class of functions taken into account. We use the orthogonal expansion approach utilizing a very specific class of rotationally invariant polynomials and possessing some other aforementioned unique properties. In the statistical literature on this subject [39, 91, 127, 210] one can find other non-parametric methods for function recovering. For example, one could apply the popular kernel estimate

$$\widetilde{f}_b(x,y) = \sum_{(x_i,y_j)\in D} \sum Z_{ij} \int_{x_i-\frac{\Delta}{2}}^{x_i+\frac{\Delta}{2}} \int_{y_j-\frac{\Delta}{2}}^{y_j+\frac{\Delta}{2}} K_b(x-u, y-v)dudv, \tag{4.35}$$

where $K_b(x,y) = b^{-2}K(b^{-1}x, b^{-1}y)$, $b > 0$ is a bandwidth parameter and $K(x,y)$ is a bivariate kernel function such that $\iint_D K(x,y)dxdy = 1$.

Using arguments developed in this chapter one can argue that the accuracy of the kernel estimate may be comparable to our method. Nevertheless, the accuracy of the kernel estimate depends not only on the choice of the bandwidth b but also on the choice of the kernel function. Indeed, the order (quantified by the number of vanishing moments) of the kernel function must be selected according to the smoothness of the unknown function. It is also known that the kernel estimate suffers from the boundary effect [39], i.e., at a boundary kernel mass falls outside the support of the function to be estimated and is lost. This effect will be further amplified by the presence of the inherent geometric error caused by the circular nature of the domain of estimated functions. It is not clear either whether the kernel estimate can have the property similar to (4.13) allowing us to construct an estimate of reduced complexity for the important class

of radial functions. Moreover, the kernel estimate is not invariant to rotations. As we have already mentioned the Zernike basis has been a popular technique in optical pattern recognition and medical imaging where one uses a few Zernike coefficients $\{\widehat{A}_{pq}\}$ to define a feature space in order to summarize complex data efficiently. Clearly the kernel estimate cannot be used for such applications.

4.5. Accuracy Analysis

In this section, we examine the components of MISE($\widehat{f}_T$) appearing in decompositions (4.32), (4.33), and (4.34). All the proofs can be found in the Appendix. In Section 4.5, we deal with the stochastic component of the error quantified by the term VAR($\widehat{f}_T$). Then the truncation error will be evaluated followed by the detailed analysis of the discretization and geometric errors. In the final subsection we summarize all the obtained results by giving the evaluation of the overall error MISE($\widehat{f}_T$). Throughout the most part of this chapter we use observation model (4.1). Nevertheless, in Section 5.5 we extend our results to the model with non-homogeneous variance.

Variance

Let us recall that the integrated variance term is given by the following formula

$$\mathrm{VAR}(\widehat{f}_T) = \sum_{p=0}^{T} \sum_{q=-p}^{p} \tau_p E|\widehat{A}_{pq} - E\widehat{A}_{pq}|^2, \tag{4.36}$$

for $p - |q|$ being even.

The bound for this term is given in the following lemma.

Lemma 4.1. *Let the observation model (4.1) be in force and let the noise process $\{\varepsilon_{ij}\}$ be uncorrelated with zero mean and variance σ^2. Then for estimate (4.28) and (4.29), we have*

$$\mathrm{VAR}(\widehat{f}_T) \le \sigma^2 \Delta^2 (T+1)(T+2)/2.$$

The proof of this bound, see [180], reveals that

$$E|\widehat{A}_{pq} - E\widehat{A}_{pq}|^2 \le \sigma^2 \Delta^2 \frac{\pi}{p+1}.$$

Using the arguments as in Section 4.5, see also the discussion below (4.65), we can show that

$$E|\widehat{A}_{pq} - E\widehat{A}_{pq}|^2 = \sigma^2\Delta^2\frac{\pi}{p+1} + O(\Delta^{2+\alpha}),$$

where $1 \leq \alpha < 3/2$ describes the geometric error.

The counterpart of Lemma 4.1, see [180], for the class of radial functions takes the following form.

Lemma 4.2. *Let the observation model (4.1) be in force and let the noise process $\{\varepsilon_{ij}\}$ be uncorrelated with zero mean and variance σ^2. Then for estimate (4.30) and (4.31), we have*

$$\mathrm{VAR}(\widehat{g}_T) \leq 2\sqrt{2}\pi\sigma^2\Delta(T+1).$$

Remark 4.5. *In Remark 4.4 we have suggested that for radial functions we should use a ring type of data grouping which utilizes all data points. It can be shown that this leads to an improved order of* $\mathrm{VAR}(\widehat{g}_T)$*, i.e., instead of* $\mathrm{VAR}(\widehat{g}_T) = O(\Delta T)$ *we have* $\mathrm{VAR}(\widehat{g}_T) = O(\Delta^2 T)$.

Truncation Error

Owing to (4.32) the truncation for $\widehat{f}_T$ is given by

$$\mathrm{TR}(\widehat{f}_T) = \sum_{p=T+1}^{\infty}\sum_{q=-p}^{p}\tau_p|A_{pq}|^2, \quad p-|q| = \text{ even} \tag{4.37}$$

whereas for $\widehat{g}_T$ by

$$\mathrm{TR}(\widehat{g}_T) = \sum_{l=T+1}^{\infty}\tau_{2l}|A_{2l}|^2. \tag{4.38}$$

It is clear that $\mathrm{TR}(\widehat{f}_T) \to 0$, $\mathrm{TR}(\widehat{g}_T) \to 0$ as $T \to \infty$, and the rate at which $\mathrm{TR}(\widehat{f}_T) \to 0$, $\mathrm{TR}(\widehat{g}_T) \to 0$ will be determined by the degree of smoothness of the functions being approximated by the Zernike functions expansion.

Let us consider the class $\mathrm{BV}(\Omega)$ of functions of bounded variation on the set Ω, i.e., functions which have finite variations in the amplitude as well as the length of the contours along which they occur. Let us recall that we say that $f \in \mathrm{BV}(\Omega)$ if the total variation

$$\mathrm{TV}(f) = \iint_{\Omega} |\bigtriangledown f(x,y)|dxdy, \tag{4.39}$$

is finite. Here the modulus of the gradient is defined as

$$|\bigtriangledown f(x,y)| = \left|\frac{\partial f(x,y)}{\partial x}\right| + \left|\frac{\partial f(x,y)}{\partial y}\right|.$$

It is important to observe that the class $\mathrm{BV}(\Omega)$ can admit discontinuous functions. In this case the derivatives are meant in the general distribution sense. For example, if $f(x,y) = \mathbf{1}_A(x,y)$ is the characteristic function of a set A having piecewise smooth boundary ∂A, then $\mathrm{TV}(f)$ is equal to the length of ∂A. It is worth noting that contrary to the one-dimensional case the bivariate function belonging to $\mathrm{BV}(\Omega)$ needs not be bounded [129, 254]. Various definitions of bounded variation functions of many variables can be found in [60, 93, 129, 254]. In particular in this book we use the concept of bounded variation for functions of two variables due to Hardy [93]. It is also worth noting that the class $\mathrm{BV}(\Omega)$ has recently been extensively used in image analysis [154, 205].

The following result gives the approximation error of the Zernike functions expansion for functions which belong to the class $\mathrm{BV}(\Omega)$. Let us recall the notation $\widetilde{f}(\rho,\theta) = f(\rho\cos\theta, \rho\sin\theta)$ for $(\rho,\theta) \in \Omega$, where $\Omega = [0,1] \times [0,2\pi]$. It is clear that the requirement $f \in \mathrm{BV}(D)$ is equivalent to $\widetilde{f} \in \mathrm{BV}(\Omega)$.

Lemma 4.3. *Let $\widetilde{f} \in \mathrm{BV}(\Omega)$ and let f be bounded. Then we have*

$$\mathrm{TR}(\widehat{f}_T) \leq \frac{d}{T+1},$$

where d depends on $\mathrm{TV}(f)$.

The proof of this important result is given in the Appendix.

The condition of bounded variation for the class of functions of the additive form introduced in (4.22) can be easily verified. In fact function (4.22) is in $\mathrm{BV}(\Omega)$ if $\varphi_s \in \mathrm{BV}([0,1])$ and $\psi_s \in \mathrm{BV}([0,2\pi])$, $s = 1, ..., L$. The ridge functions in (4.27) are in $\mathrm{BV}(\Omega)$ if each $g_l(t)$ belongs to $\mathrm{BV}([-1,1])$.

Regarding the case of radial functions let us assume that $g(\rho^2)$ is in $\mathrm{BV}([0,1])$. The following lemma, see [180], gives a bound for $\mathrm{TR}(\widehat{g}_T)$.

Lemma 4.4. *Let $g \in \mathrm{BV}([0,1])$. Then we have*

$$\mathrm{TR}(\widehat{g}_T) \leq \frac{8(\overline{g} + \mathrm{TV}(g))}{T+1/2},$$

where $\mathrm{TV}(g)$ is the total variation of $g(t)$ on $[0,1]$ and $\overline{g} = \max_{t\in[0,1]} |g(t)|$.

Lemmas 4.3 and 4.4 hold for functions of bounded variation. This class allows functions that are discontinuous, which is often the case in image processing applications. Let us consider a class of Lipschitz continuous functions, i.e.,

$$|f(x,y) - f(x',y')| \leq L\{|x-x'|^\alpha + |y-y'|^\alpha\}, \quad 0 < \alpha \leq 1$$

for (x,y), $(x',y') \in D$ and a positive constant L.

To evaluate the truncation error for this case one can use the known results [224] concerning the best polynomial approximation of Lipschitz continuous functions. This yields the following bound

$$\mathrm{TR}(\widehat{f}_T) \le \frac{a_1}{(T+1)^{2\alpha-1}}, \tag{4.40}$$

where $\alpha > 1/2$ and the constant a_1 depends on L. Similarly if one assumes that all partial derivatives $\dfrac{\partial^{i+j} f(x,y)}{\partial x^i \partial y^j}$, $i+j \le r$ exist and that $\dfrac{\partial^r f(x,y)}{\partial x^i \partial y^j}$, $i+j=r$ are Lipschitz (with $\alpha = 1$) then

$$\mathrm{TR}(\widehat{f}_T) \le \frac{a_2}{(T+1)^{2r+1}}, \tag{4.41}$$

where the constant a_2 depends on the L_2-norms of $\dfrac{\partial^{i+j} f(x,y)}{\partial x^i \partial y^j}$, $i+j=r$ and the Lipschitz constant L.

The bounds in (4.40) and (4.41) can be improved in some special cases. For example, result (4.40) can be improved in the case of radial functions as it is shown in the following lemma, see [180] for the proof.

Lemma 4.5. *Let $g^{(1)} \in L_2([0,1])$. Then we have*

$$\mathrm{TR}(\widehat{g}_T) \le \frac{\beta}{(T+1)^2},$$

where

$$\beta = \pi \int_0^1 t(1-t)|g^{(1)}(t)|^2 dt.$$

Hence we have the rate $O(T^{-2})$ which should be compared with $O(T^{-1})$ in (4.40) for $\alpha = 1$.

The rate in Lemma 4.5 can also be extended to the class of additive functions defined in (4.25), i.e., the functions of the following form

$$\widetilde{f}(\rho,\theta) = \varphi(\rho^2) + \psi(\theta). \tag{4.42}$$

Formula (4.26) gives the Zernike coefficient for the additive functions. The first term in (4.26) is identical to the Zernike coefficient of radial functions. As for the second term it can be proved [180] that it is exponentially small. Hence the angular component in (4.42) plays a smaller role than the radial one in the overall approximation error. This is formalized in the following lemma.

Lemma 4.6. *Let $f(x,y)$ be of the additive form (4.42). Let $\psi \in L_2([0,2\pi])$ and let $\varphi^{(1)} \in L_2([0,1])$. Then we have*

$$\mathrm{TR}(\widehat{f}_T) \leq \frac{c_1}{(T+1)^2} + c_2 \exp(-c_3 T),$$

for some positive constants c_1, c_2, and c_3.

It should be noted that the constant c_1 can be selected as 4β, where β is given in Lemma 4.5. To get further insight into the order of the approximation error let us consider a specific class of functions defined on D.

Example 4.1. *Let $f(x,y)$ be radial, denote this by $f(x,y) = g(x^2+y^2)$. First, consider the function*

$$g_1(t) = (2t)^\mu, \quad 0 \leq t \leq 1, \quad \mu > 0. \tag{4.43}$$

It is clear that $g_1 \in \mathrm{BV}([0,1])$ and also satisfies the condition of Lemma 4.5 if $\mu > 1/2$.

The discussion in Section 4.3 shows, see (4.13), that for radial functions we need to evaluate the following integral

$$\int_0^1 g_1(t) P_k(2t-1) t dt, \tag{4.44}$$

where $P_k(x)$ is the k-th order Legendre polynomial.

Using the result in [10] we can show that (4.44) is of order $O\left((k+1)^{-2(\mu+1)}\right)$. This being applied in (4.38) yields

$$\mathrm{TR}(\widehat{g}_T) = O((T+1)^{-2(2\mu+1)}). \tag{4.45}$$

In the same way we can argue that for the function $g_2(t) = (2(1-t))^\mu$, $0 \leq t \leq 1$, $\mu > 0$, the rate in (4.45) holds. It should be noted that both $g_1(t)$ and $g_2(t)$ are monotonic on $[0,1]$.

Let us now consider the function $g_3(t) = |2(t-1/2)|^\mu$, $0 \leq t \leq 1$, $\mu > 0$ which is not monotonic but convex on $[0,1]$. Again using [10] we can obtain that the integral in (4.44) is of order $O\left((k+1)^{-(\mu+3/2)}\right)$ and consequently we have

$$\mathrm{TR}(\widehat{g}_T) = O((T+1)^{-(2\mu+1)}).$$

This is a twice slower rate than in (4.45). Nevertheless, the above rates are faster than predicted by our bounds. Hence we can conjecture that for a large class of image functions, e.g., being convex, monotonic functions, we can improve our rates.

The following characteristic represents the class of discontinuous piecewise constant functions

$$g_4(t) = a\mathbf{1}_{\{t\leq\gamma\}}(t) + b\mathbf{1}_{\{\gamma<t\}}(t), \tag{4.46}$$

where $0 < \gamma < 1$ *and, without loss of generality,* $a > b$*. Note that* $g_4 \in \mathrm{BV}([0,1])$ *with* $\mathrm{TV}(g_4) = a - b$*. A tedious but otherwise straightforward algebra making use of some properties of the Legendre polynomials [206] yields the following bound*

$$\left|\int_0^1 g_4(t)P_k(2t-1)tdt\right| \leq \frac{\mathrm{TV}(g_4)}{\sqrt{\pi}(2k+1)\sqrt{k+2}}.$$

This leads to

$$\mathrm{TR}(\widehat{g}_T) \leq \frac{8(a-b)^2}{T+1}.$$

This is consistent with the general result given in Lemma 4.4.

Let us finally consider the case of an unbounded function and neither being in $\mathrm{BV}([0,1])$*. Hence let*

$$g_5(t) = \frac{1}{|2(t-1/2)|^\mu}, \quad 0 \leq t \leq 1.$$

It is clear that $g_5 \in L_2([0,1])$ *if only* $0 < \mu < 1/2$*. Some algebra shows that*

$$\left|\int_0^1 g_5(t)P_k(2t-1)tdt\right| = O((k+1)^{3/2-\mu}).$$

Hence we obtain

$$\mathrm{TR}(\widehat{g}_T) = O((T+1)^{-(1-2\mu)}), \quad 0 < \mu < 1/2.$$

This is clearly a very slow rate of convergence.

Geometric and Discretization Errors

In order to explain the nature of the geometric error let us consider the indicator function of D, i.e., $f(x,y) = \mathbf{1}_D(x,y)$. Then we obtain

$$E\widehat{A}_{pq} - A_{pq} = \sum_{(x_i,y_j)\in D}\sum w_{pq}(x_i,y_j) - \iint_D V^*_{pq}(x,y)dxdy.$$

Noting that $\iint_D V^*_{pq}(x,y)dxdy = \pi$ if only $p = q = 0$ and $\iint_D V^*_{pq}(x,y)dxdy = 0$ otherwise, it suffices to consider the case $p = q = 0$, i.e.,

$$E\widehat{A}_{00} - A_{00} = \sum_{(x_i,y_j)\in D}\sum w_{00}(x_i,y_j) - \pi. \tag{4.47}$$

The above term is not equal to zero due to the fact that if the center of a lattice square falls inside the border of the unit disk D, this square is used in the computation of the estimate $\widehat{A}_{pq}$ otherwise, the square is discarded. Therefore, the area used for computing $\widehat{A}_{pq}$ is not equal to the area of the unit disk. Fig. 3.4 shows the union of the squares whose centers fall inside the unit circle. Note that some squares are not entirely inside the circle; on the other hand, some parts of the circle are not covered by the squares. This boundary region is shown in Fig. 4.3 where only the geometric error within the unit disk is depicted.

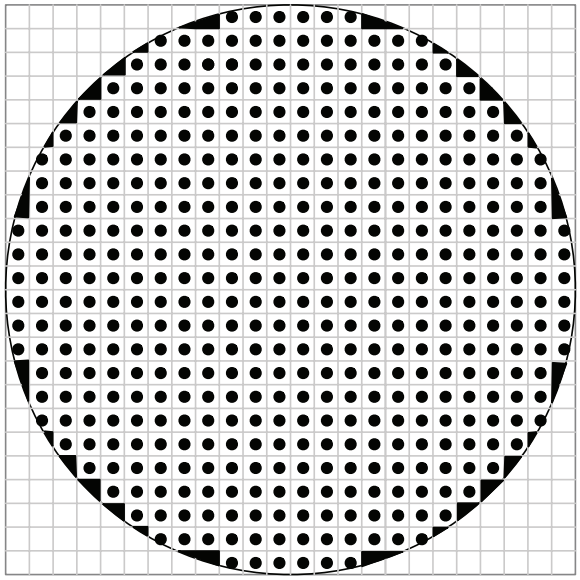


Figure 4.3: The boundary region defining the geometric error

Recalling that $V_{00}^{*}(x,y) = 1$ we can rewrite (4.47) as follows

$$E\widehat{A}_{00} - A_{00} = \Delta^2 K(\Delta) - \pi, \tag{4.48}$$

where

$$K(\Delta) = \sum_{(x_i,y_j)\in D}\sum 1$$

denotes the number of the points $\{(x_i, y_j) : 1 \leq i, j \leq n\}$ inside the unit circle.

As we shall see formula (4.48) fully describes the geometric error. Hence let us define, similarly as in Chapter 3, the following number

$$G(\Delta) = \Delta^2 K(\Delta) - \pi. \tag{4.49}$$

Since our theory models the performance of function recovery on the grid which becomes increasingly fine therefore it is crucial to know the size of $G(\Delta)$, i.e., how fast $G(\Delta)$ tends to zero as $\Delta \to 0$.

As we have already noticed the quantity $G(\Delta)$ has been extensively examined in the analytic number theory with the relation to the so-called lattice points of a circle problem due originally to Gauss [87, 104–106, 110, 235].

We have observed that we have $G(\Delta) = O(\Delta^{\alpha})$, $1 \leq \alpha < 3/2$ where $\alpha = 3/2$ is impossible and the best result so far due to Huxley is the following bound [106]

$$G(\Delta) = O(\Delta^{285/208}). \tag{4.50}$$

It has been conjectured that the value of α can be arbitrary close to $\alpha = 3/2$ This still remains an open problem in the analytic number theory, see [87,105,106,235].

Hence the conjectured rate is of order

$$G(\Delta) = O(\Delta^{3/2-\varepsilon}). \tag{4.51}$$

A numerical study has been conducted to evaluate $G(\Delta)$ for a large range of Δ. In Fig. 4.4 we depict $|G(\Delta)|$ as a function of $n = 2/\Delta$ for $10 \leq n \leq 128$, i.e., $0.015625 \leq \Delta \leq 0.2$.

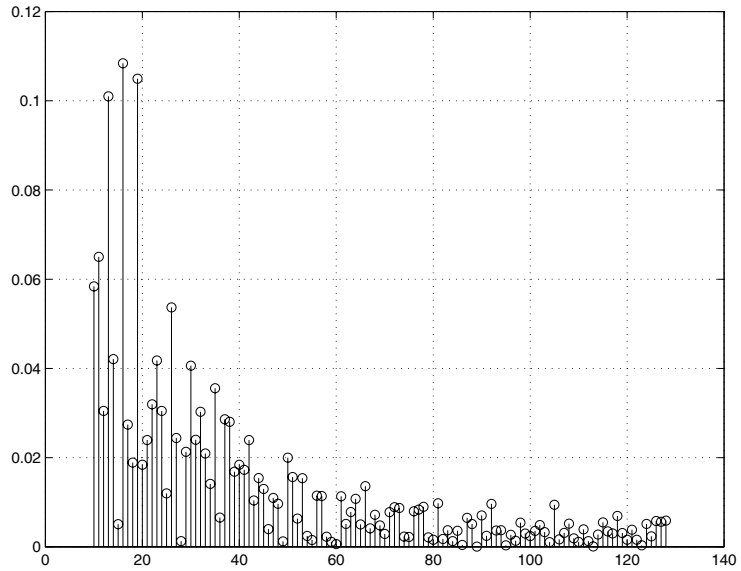


Figure 4.4: The geometric error $|G(\Delta)|$ versus n, $10 \leq n \leq 128, n = 2/\Delta$

The oscillatory behavior of $G(\Delta)$ can be observed with the error taking both positive and negative values. However, the positive values (around 78%) are more

frequent than the negative ones. Since the number $G(\Delta)$ influences the accuracy of the estimate $\widehat{A}_{pq}$, therefore $\widehat{A}_{pq}$ is more likely to be positively biased. We refer to [55] for similar numerical results concerning the evaluation of $G(\Delta)$.

Let us now consider the term $\mathrm{DG}(\widehat{f}_T)$ in (4.6), i.e.,

$$\mathrm{DG}(\widehat{f}_T) = \sum_{p=0}^{T} \sum_{q=-p}^{p} \tau_p |E\widehat{A}_{pq} - A_{pq}|^2.$$

To evaluate this term it suffices to consider the difference $E\widehat{A}_{pq} - A_{pq}$. First it is an important observation that $E\widehat{A}_{pq} - A_{pq}$ can be decomposed into two separate components describing the discretization and geometric errors, i.e.,

$$E\widehat{A}_{pq} - A_{pq} = D_{pq}(\widehat{f}_T) + G_{pq}(\widehat{f}_T), \tag{4.52}$$

where

$$D_{pq}(\widehat{f}_T) = \sum_{(x_i,y_j)\in D} \sum \iint_{p_{ij}\cap D} (f(x_i, y_j) - f(x, y))\, V_{pq}^*(x, y)dxdy \tag{4.53}$$

and

$$G_{pq}(\widehat{f}_T) = -\iint_{\Theta(\Delta)} f(x, y)V_{pq}^*(x, y)dxdy. \tag{4.54}$$

Here p_{ij} denotes the (i, j) lattice square centered at (x_i, y_j) and $\Theta(\Delta)$ is the region being the intersection of D with the union of those grid squares whose centers (x_i, y_j) fall outside the circle, see Fig. 4.3.

The magnitude of the discretization error is given in the following lemma, see [180] for the proof.

Lemma 4.7. *Let $f \in \mathrm{BV}(D)$ and let f be bounded. Then for any admissible pair (p, q) we have*

$$|D_{pq}(\widehat{f}_T)| \le \left\{ \frac{\overline{f}\pi}{p+1} \mathrm{TV}(f) \right\}^{1/2} \Delta,$$

where $\mathrm{TV}(f)$ is the total variation of f and $\overline{f} = \max_{x,y} |f(x, y)|$.

The magnitude of the discretization error can be further reduced assuming a bit stronger conditions on $f(x, y)$. Hence if $f(x, y)$ is Lipschitz of order 1 then by a simple modification of the proof of Lemma 4.7, we obtain

$$\left| D_{pq}(\widehat{f}_T) \right| \le c \frac{\Delta^{3/2}}{(p+1)^{1/2}} \tag{4.55}$$

for some constant c.

The discretization error for the estimate $\widehat{g}_T(\rho^2)$ in (4.30) is described in the following lemma. See [180] for the proof.

Lemma 4.8.

(a) *Let $g \in \mathrm{BV}([0,1])$. Then*

$$|E\widehat{A}_{2k} - A_{2k}| \le 2\pi\{\mathrm{TV}(g)\overline{g}\}^{1/2}\left(\frac{\Delta}{2k+1}\right)^{1/2}.$$

(b) *Let $g^{(1)} \in L_2([0,1])$. Then*

$$|E\widehat{A}_{2k} - A_{2k}| \le \left\{2\int_0^1 \left|g^{(1)}(t)\right|^2 dt\right\}^{1/2} \frac{\Delta}{(2k+1)^{1/2}}.$$

Let us now turn to the geometric error expressed by formula (4.54). It is important to observe that the area of the region $\Theta(\Delta)$ in (4.54) is bounded by the error of the lattice approximation of D defined in (4.49), i.e., we have

$$\mathrm{Area}\,(\Theta(\Delta)) \le |G(\Delta)|. \tag{4.56}$$

Using the arguments as in Section 3.4 we obtain the following result.

Lemma 4.9. *Let $f \in L_2(D)$ be a bounded function on D. Then for any admissible pair (p,q) we have*

$$|G_{pq}(\widehat{f}_T)| \le \overline{f}|G(\Delta)|^{1/2}\min\left(\left\{\frac{\pi}{p+1}\right\}^{\frac{1}{2}}, |G(\Delta)|^{1/2}\right).$$

Hence the geometric error depends on the lattice approximation factor $G(\Delta)$ introduced in (4.49). Using the best known bound so far on the decay of $G(\Delta)$ due to Huxley [106] we obtain

$$|G_{pq}(\widehat{f}_T)| = O\left(\Delta^{285/416}\min(p^{-1/2}, \Delta^{285/416})\right). \tag{4.57}$$

Lemma 4.7 and Lemma 4.9 yield the following result concerning the discretization and geometric term $\mathrm{DG}(\widehat{f}_T)$.

Lemma 4.10. *Let $f \in \mathrm{BV}(D)$ and let f be bounded. Then we have*

$$\begin{aligned}\mathrm{DG}(\widehat{f}_T) \le{}& (T+1)(T+2)\overline{f}\,\mathrm{TV}(f)\Delta^2 \\ &+ \overline{f}^2|G(\Delta)|\min\left((T+1)(T+2), \frac{1}{\pi}T(T+1)(T+2)|G(\Delta)|\right).\end{aligned}$$

Thus with the best known result [106] on the lattice approximation of circle we obtain

$$\mathrm{DG}(\widehat{f}_T) = O(T^2\Delta^2) + O\left(\min(T^2\Delta^{285/208}, T^3\Delta^{285/104})\right). \tag{4.58}$$

The following lemma summarizes the analogous results for radial functions. This is implied by Lemma 4.8 and formula (4.34).

Lemma 4.11.

(a) *Let the conditions of Lemma 4.8 (a) hold. Then we have*

$$D(\widehat{g}_T) \leq c_1\Delta(T+1).$$

(b) *Let the conditions of Lemma 4.8 (b) hold. Then we have*

$$D(\widehat{g}_T) \leq c_2\Delta^2(T+1).$$

Using the bounds obtained in Lemma 4.8 the constants c_1, c_2 are given by $c_1 = 4\pi V(g)\overline{g}$, and $c_2 = 2 \parallel g^{(1)} \parallel^2 /\pi$.

MISE

Taking into consideration the aforementioned results one can now easily evaluate the overall estimation error $\mathrm{MISE}(\widehat{f}_T)$. In fact combining Lemmas 4.1, 4.3, and 4.10 we obtain:

Theorem 4.1. *Let $f \in L_2(D) \cap \mathrm{BV}(D)$ and f be bounded. Then we have*

$$\begin{aligned}\mathrm{MISE}(\widehat{f}_T) \leq \sigma^2\Delta^2(T+1)(T+2)/2 + (T+1)(T+2)\overline{f}\,\mathrm{TV}(f)\Delta^2 + \frac{d}{T+1} \\ + \overline{f}^{\,2}|G(\Delta)| \min\left((T+1)(T+2), \pi^{-1}T(T+1)(T+2)|G(\Delta)|\right),\end{aligned} \tag{4.59}$$

for some positive constant d.

Remark 4.6. *Formula (4.59) shows an apparent trade-off between the variance, the discretization, geometric errors, and the truncation error. Thus using the best known bound on the decay of $G(\Delta)$ due to Huxley [106] we obtain*

$$\mathrm{MISE}(\widehat{f}_T) = O(T^2\Delta^2) + O(\min(T^2\Delta^{285/208}, T^3\Delta^{285/104})) + \frac{d}{T+1}. \tag{4.60}$$

It is important to note that although the overall error is dominated asymptotically by the variance term and the truncation error, we cannot ignore the geometric error in practical application of the Zernike moments.

Formula (4.60) leads to the following asymptotic rate of convergence for $\mathrm{MISE}(\widehat{f}_T)$.

Corollary 4.1. *Under the conditions of Theorem 4.1 we have*

$$\mathrm{MISE}(\widehat{f}_{T^*}) = O(\Delta^{2/3}), \tag{4.61}$$

with the truncation parameter selected as $T^* = a\Delta^{-2/3}$.

We can conjecture that (4.61) represents the best possible rate of any linear nonparametric recovering technique for the problem of estimating an image of the class $\mathrm{BV}(D)$.

Note also that for the selected $T^* = a\Delta^{-2/3}$ the contribution from the geometric error is of order $O(T^{*3}\Delta^{285/104}) = O(\Delta^{77/104})$. This should be compared with the rate of the decay of the variance term and the truncation error being $O(\Delta^{2/3})$.

The $\mathrm{MISE}(\widehat{f}_T)$ for images satisfying the condition leading to the approximation error in (4.41) is of the following order

$$\mathrm{MISE}(\widehat{f}_T) = O(\Delta^{\frac{2(2r+1)}{2r+3}}) \tag{4.62}$$

with the truncation parameter selected as

$$T^* = a\Delta^{-\frac{2}{2r+3}}.$$

Hence we can conjecture that for smooth functions, the optimal rate of convergence is of the order

$$\mathrm{MISE}(\widehat{f}_T) = O\left(\Delta^{2-\varepsilon}\right).$$

The rate given in (4.62) can be improved if some structural properties of $f(x, y)$ are exploited. In fact Lemma 4.6 shows that $\mathrm{TR}(\widehat{f}_T)$ for the class of additive functions is of order $O(T^{-2})$. By this, Lemma 4.2 and the result in (4.58) we obtain

$$\mathrm{MISE}(\widehat{f}_T) = O\left(\Delta\right)$$

with T selected as $T^* = a\Delta^{-1/2}$.

Let us now turn to the case of radial functions, i.e., let us consider the accuracy of the estimate $\widehat{g}_T$. Recalling Lemmas 4.2, 4.4, and 4.8 we obtain:

Theorem 4.2.

(a) *Let* $g \in \mathrm{BV}([0,1])$. *Then we have*

$$\mathrm{MISE}(\widehat{g}_T) \le 2\sqrt{2}\pi\sigma^2(T+1)\Delta + 4\pi V(g)\overline{g}(T+1)\Delta + 8(V(g)+\overline{g})(T+1/2)^{-1}.$$

(b) *If in turn* $g^{(1)} \in L_2([0,1])$ *then*

$$\begin{aligned} \mathrm{MISE}(\widehat{g}_T) \quad &\leq \quad 2\sqrt{2}\pi\sigma^2(T+1)\Delta + 2\pi^{-1} \parallel g^{(1)} \parallel^2 (T+1)\Delta^2 \\ &\quad + 4^{-1}\pi \parallel g^{(1)} \parallel^2 (T+1)^{-2}, \end{aligned} \tag{4.63}$$

where $\parallel g^{(1)} \parallel$ *is the* $L_2([0,1])$ *norm of* $g^{(1)}$.

The last term in (4.63) is derived from the bound in Lemma 4.5 by observing that $\beta \leq 4^{-1}\pi \parallel g^{(1)} \parallel^2$. Direct minimization of the bounds in Theorem 4.2 yields the following result concerning the rate of convergence.

Corollary 4.2. *Under the condition of Theorem 4.2 (a) we have*

$$\mathrm{MISE}(\widehat{g}_T) = O(\Delta^{1/2})$$

with the truncation parameter selected as $T^* = a\Delta^{-1/2}$.

In turn under the condition of Theorem 4.2 (b) we obtain

$$\mathrm{MISE}(\widehat{g}_T) = O(\Delta^{2/3})$$

with the truncation parameter selected as $T^* = a\Delta^{-1/3}$.

The above rates can be improved if one applies the ring data grouping strategy suggested in Remark 4.4. In fact using Remark 4.5 we obtain the following result.

Corollary 4.3. *Under the condition of Theorem 4.2 (a) we have*

$$\mathrm{MISE}(\widehat{g}_T) = O\left(\Delta\right)$$

with the truncation parameter selected as $T^* = a\Delta^{-1}$.

In turn under the condition of Theorem 4.2 (b) we obtain

$$\mathrm{MISE}(\widehat{g}_T) = O(\Delta^{4/3})$$

with the truncation parameter selected as $T^* = a\Delta^{-2/3}$.

Example 4.2. *To illustrate the above results we plot in Fig 4.5 the bound in (4.59). We use* $\sigma^2 = 0.1$ *with the signal to noise ratio equal 10, hence* $\overline{f} = 1$, $V(f) = 1$, *and also* $d = 1$. *We take* $\Delta = 2/n$ *with* $n = 990$. *In this case our numerical studies (partially illustrated in Fig. 4.5) show that* $G(\Delta)$ *is of order 0.000323477. It is seen that the optimal value of the truncation parameter is given by* $T^* = 10$. *In the same plot we depict bound (4.59) with the bias term* $d/(T+1)$ *replaced by* $d/(T+1)^3$. *This corresponds to functions with all partial derivatives of order one satisfying the Lipschitz condition, see (4.41). It can be observed that now* $T^* = 4$.

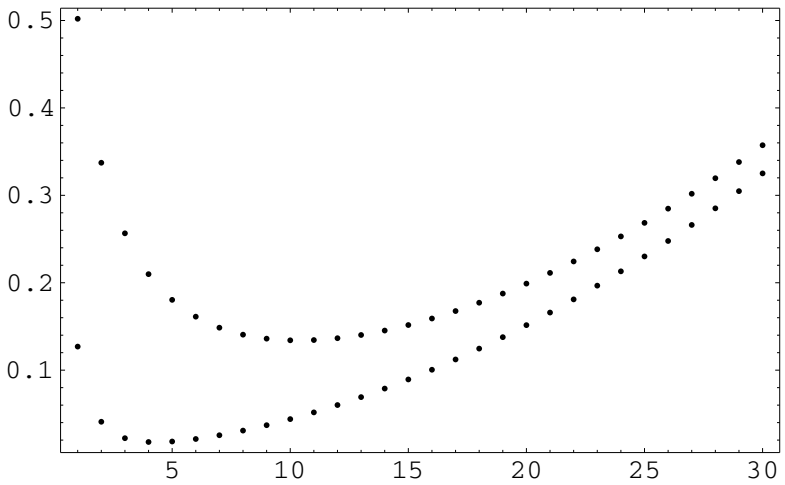


Figure 4.5: MISE($\widehat{f}_T$) versus T, $1 \leq T \leq 30$ for two different classes of functions

Remark 4.7. *In practice it is of great importance to know how the truncation parameter T should be chosen for a given, finite sample size. For some general results on data driven selection of smoothing and regularization parameters in nonparametric curve estimation we refer to [39] and the references cited therein. In our context one could use the well-studied cross-validation method seeking T_{CV} which minimizes*

$$CV(T) = \Delta^2 \sum_{i=1}^{n} \sum_{j=1}^{n} \left\{ Z_{ij} - \widehat{f}_{T,ij}(x_i, y_j) \right\}^2,$$

where $\widehat{f}_{T,ij}(x,y)$ is an estimate of the same type as $\widehat{f}_T(x,y)$ except that it is computed without data value $((x_i, y_j), Z_{ij})$. It should be noted that in the definition of $\{w_{pq}(x_r, y_s)\}$ in (4.28) we have to integrate over two pixels in the case of pixels adjacent to the pixel (x_i, y_j). The $CV(T)$ criterion corresponds to squared error

$$SE(T) = \Delta^2 \sum_{i=1}^{n} \sum_{j=1}^{n} \left\{ \widehat{f}_T(x_i, y_j) - f(x_i, y_j) \right\}^2,$$

which measures the discrepancy between $\widehat{f}_T(x,y)$ and $f(x,y)$ on the grid points. Let T_{SE} be the truncation value minimizing $SE(T)$. It is an interesting problem to assess the accuracy of T_{CV} as an estimate of T_{SE}. Using the results of this

chapter and techniques developed in [92] we can show that

$$\frac{T_{CV}}{T_{SE}} = 1 + O(\Delta^{1/3}),$$

in probability.

4.6. Generalizations

Thus far we have considered rather classical observation model defined in (4.1). In a number of applications of image analysis, one is confronted with the so-called signal dependent noise model, see [112] and references cited therein. Hence (4.1) is replaced by

$$Z_{ij} = f(x_i, y_j) + \sqrt{s(x_i, y_j)}\varepsilon_{ij}, \quad (x_i, y_j) \in D, \quad 1 \le i, j \le n, \tag{4.64}$$

where $s(x, y)$ is a non-negative function representing the inhomogeneity of the noise process. Examples include the film grain noise, where $s(x, y) = f^{2/3}(x, y)$ and the multiplicative noise model where $s(x, y) = f^2(x, y)$. It is clear that this type of noise model is going merely to influence the variance of the estimate $\widehat{f}_T(x, y)$. Hence the term $\mathrm{VAR}(\widehat{f}_T)$ in decomposition (4.33) has to be examined. Proceeding as in the proof of Lemma 4.1 we can obtain

$$\mathrm{var}(\widehat{A}_{pq}) = \sigma^2 \sum_{(x_i,y_j)\in D} \sum \left| \iint_{p_{ij}\cap D} V^*_{pq}(x, y)dxdy \right|^2 s(x_i, y_j). \tag{4.65}$$

Using arguments as in (4.52), i.e., observing that

$$\begin{aligned} \iint_D |V_{pq}(x, y)|^2 \, s(x, y)dxdy &= \sum_{(x_i,y_j)\in D} \sum \iint_{p_{ij}\cap D} |V_{pq}(x, y)|^2 \, s(x, y)dxdy \\ &\quad + \iint_{\Theta(\Delta)} |V_{pq}(x, y)|^2 \, s(x, y)dxdy \end{aligned}$$

we can approximate the sum in (4.65) as follows

$$\Delta^2 \left(\iint_D |V_{pq}(x, y)|^2 \, s(x, y)dxdy + \xi_{pq}(\Delta) \right), \tag{4.66}$$

where $\xi_{pq}(\Delta)$ is the error term of the approximation. Clearly $\xi_{pq}(\Delta)$ is of order $G(\Delta)$. The dependence of $\mathrm{VAR}(\widehat{f}_T)$ on $G(\Delta)$, however, is weaker than in the bias term as $G(\Delta)$ contributes in (4.66) to the second order effect.

The results of this chapter can also be extended to other directions. For instance one can consider the 3-D counterpart of the Zernike functions which finds applications in 3-D computer vision, computer graphics, and crystallography. In fact the following

$$V_{pq}^{m}(x,y,z) = R_{pq}(\rho)Y_{qm}(\theta,\vartheta)$$

is a complete and orthogonal set of functions in $L_2(B)$, where B is the 3-D unit ball. Here ρ is the radial component and θ, ϑ are angular variables. In the above formula $Y_{qm}(\theta,\vartheta)$ is the spherical harmonic defined as

$$Y_{qm}(\theta,\vartheta) = C_{qm}P_{q}^{m}(\cos\theta)e^{jm\vartheta},$$

where C_{qm} is the normalizing constant, and $P_{q}^{m}(x)$ denotes an associated Legendre polynomial [217]. Spherical harmonics form a Fourier basis on the sphere analogously as $\{e^{jm\theta}\}$ does on a circle.

The problem of lattice points which is critical for our developments has also been studied for B. Hence let $K(\Delta)$ have the meaning as in (4.49), i.e., it is the number of lattice cubes falling into B. Then it is known that [12, 87, 105, 235]

$$\Delta^3 K(\Delta) = 4\pi/3 + O(\Delta^2).$$

It is a rather surprising result that the geometric error in the 3-D case is of lower order than the one for the 2-D circle.

4.7. Conclusions

In this chapter we have presented a basic theory of function recovery defined on the circular domain. The estimate based on a class of invariant orthogonal functions has been proposed and its statistical properties have been evaluated. Various bounds for the mean integrated squared error have been demonstrated. These bounds have been obtained by decomposing the error into four non-interfering terms that reflect different errors such as the amount of noise in the data, the discretization error, the geometric error, and the approximation term. In fact the estimate $\widehat{A}_{pq}$ of the Zernike moment A_{pq} is decomposed as follows

$$\widehat{A}_{pq} = A_{pq} + (\widehat{A}_{pq} - E\widehat{A}_{pq}) + (E\widehat{A}_{pq} - A_{pq}), \tag{4.67}$$

where the first term in the brackets represents the stochastic component of $\widehat{A}_{pq}$, whereas the second term in the brackets is the bias term of $\widehat{A}_{pq}$. We show that $E\widehat{A}_{pq} - A_{pq}$ can be decomposed further into the geometric and numerical errors. The stochastic term is $O(\Delta)$ (in probability), whereas the numerical error is

$O(\Delta^{\beta})$, where $\beta \geq 1$ depends on the smoothness of $f(x, y)$. The geometric error on the other hand can never exceed $\Delta^{3/2}$. The best known result on the circle lattice problem [87, 104, 105] yields the geometric error of order $O\left(\Delta^{285/208}\right)$. Although our results are based on upper bounds, we believe that the orders of convergence rates are correct and that the geometric error is an inherent factor of any function reconstruction problem defined on non-rectangular domain.

In fact, we can define orthogonal polynomial basis on general convex domains [126]. In particular, the elliptical and triangular (or convex polygons in general) domains (see Fig. 4.6) are important in computer vision, graphics, and medical imaging. Our theory can be extended to these cases by noting that there exist corresponding lattice point approximations. In the case of an ellipse $\mathcal{E}$ the lattice points approximation reads as

$$\Delta^2 \sum_{(x_i,y_j)\in\mathcal{E}} 1 = |\mathcal{E}| + O(\Delta^{15/11}),$$

whereas for a triangle $\mathcal{T}$ we have

$$\Delta^2 \sum_{(x_i,y_j)\in\mathcal{T}} 1 = |\mathcal{T}| + O(\Delta),$$

where $|\mathcal{E}|$ and $|\mathcal{T}|$ stand for the area of $\mathcal{E}$ and $\mathcal{T}$, respectively. Lattice approximation results concerning other convex domains in R^2 can be found in [104] and [105].

Then, using the above results and techniques developed in this chapter we can show that the reconstruction error $MISE(\widehat{f}_T^{\mathcal{E}})$ for the ellipse is of order $O(\Delta^{2/3})$, i.e., it is identical to that for the circular domain. This is due to the fact that the geometric error is of the second order importance compared to the variance term. This is in striking contrast with the triangular domain where the error $MISE(\widehat{f}_T^{\mathcal{T}})$ is of order $O(\Delta^{1/2})$. Indeed, the geometric error in the decomposition of $MISE(\widehat{f}_T^{\mathcal{T}})$ is $O(T^3\Delta^2)$ which dominates the variance term being $O(T^2\Delta^2)$.

Generally, we can conjecture that for convex domains with smooth boundaries (like ellipse) the geometric error is not the dominant factor, whereas for the convex domains with non-smooth boundaries (like polygons) the reconstruction error is solely dominated by the geometric error.

We have also observed that the geometric error can be eliminated for the case of radial functions for which the Zernike coefficients take a particularly simple form (formula (4.13)) allowing one to define the simplified estimate $\widehat{g}_T$, see (4.30). In practice, however, it is difficult to verify whether the observed data in (4.1) represent a radial function. It is interesting to note, however, that property (4.13) allows us to form a simple statistical test to verify the radiality of the

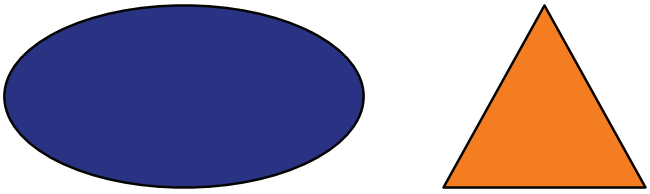


Figure 4.6: Smooth (ellipse) and non-smooth (triangle) convex domains

underlying image. In Chapter 6 we develop formal testing procedures for verifying the radiality of an image function observed in the presence of noise. On the other hand, in Chapter 5 we develop an efficient algorithm for computing the Zernike moments which entirely eliminates the geometric error.

Our estimate based on the radial polynomials, being global functions in nature, suffers degradation in accuracy due to the geometric error. To alleviate the error one could use local multiresolution/wavelet expansions [111, 154] with possible different spatial scales. The issue of invariance, however, remains still open as standard wavelet expansions do not have this desirable property. This topic is left for future studies.

In summary, the following results have been obtained in this chapter:

- Basic approximation properties of the Zernike orthogonal radial polynomials have been established.
- The behavior of Zernike moments for images with various symmetry and structural conditions has been examined. This includes radial images and images of the additive form.
- An image reconstruction formula from the Zernike orthogonal radial polynomials has been introduced.
- The detailed error analysis of the reconstruction algorithm has been carried out. This includes the noise diminishing property and the imprecision of the Zernike moment calculation due to the geometric and discretization errors.
- The fundamental bound for the truncation reconstruction error has been established.
- The rates for the mean integrated squared error have been derived.
- The problem of data-driven selection of the optimal number of moments has been discussed.

4.8. Appendix

For further considerations we shall need the following two auxiliary results.

Lemma 4.12. *Let $h(x,y)$ satisfy the following conditions:*

$$\begin{aligned} h(x,y) &\geq 0, \\ h(x,y)-h(x,\eta) &\geq 0, \text{ for all } y \geq \eta, \\ h(x,y)-h(\xi,y) &\geq 0, \text{ for all } x \geq \xi, \\ h(x,y)-h(x,\eta)-h(\xi,y)+h(\xi,\eta) &\geq 0, \text{ for all } x \geq \xi \text{ and } y \geq \eta. \end{aligned} \tag{4.68}$$

(a) *Any function $f(x,y)$ whose variation in $[a,A]\times[b,B]$ is finite may be expressed in the form*

$$f(x,y) = f_1(x,y) - f_2(x,y),$$

where f_1 and f_2 satisfy conditions (4.68).

(b) *Let $g(x,y)$ be a finite and integrable function on $[a,A]\times[b,B]$, and let $f(x,y)$ satisfy conditions (4.68). Then*

$$\int_a^A\int_b^B f(x,y)g(x,y)dxdy = f(A-,B-)\int_\xi^A\int_\eta^B g(x,y)dxdy,$$

where $a \leq \xi \leq A$, $b \leq \eta \leq B$.

These results were proved originally in [93] and are not easily accessible in standard books on multivariate calculus. Note that the result in (a) represents a counterpart of the classical decomposition of a univariate function of bounded variation as a difference of two non-decreasing and positive functions. On the other hand, the part (b) is a generalization of the second mean-value theorem to the case of two dimensional functions.

Proof of Lemma 4.3. Let $\mathrm{TV}(f)$ denote the total variation of $f(x,y)$ and let $\overline{f} = \max_{x,y}|f(x,y)|$. By virtue of Lemma 4.12 (a) and (4.10) we can write

$$A_{pq} = \int_0^{2\pi}\int_0^1 \widetilde{f}_1(\rho,\theta)R_{pq}(\rho)e^{-jq\theta}\rho d\rho d\theta - \int_0^{2\pi}\int_0^1 \widetilde{f}_2(\rho,\theta)R_{pq}(\rho)e^{-jq\theta}\rho d\rho d\theta, \tag{4.69}$$

where $\widetilde{f}_1(\rho,\theta)$ and $\widetilde{f}_2(\rho,\theta)$ are functions defined over the rectangle $[0,2\pi]\times[0,1]$ satisfying conditions (4.68). A quick inspection of the decomposition given in [93] reveals that $0 \leq \widetilde{f}_1(\rho,\theta) \leq 2\overline{f} + 3\,\mathrm{TV}(f)$, $0 \leq \widetilde{f}_2(\rho,\theta) \leq 2\overline{f} + 3\,\mathrm{TV}(f)$. Using the

bound for $R_{pq}(\rho)$ established in (4.7) one can observe that $g(\rho,\theta) = R_{pq}(\rho)e^{-jq\theta}\rho$ is a finite and integrable function on $[0,2\pi]\times[0,1]$. Thus by Lemma 4.12 (b) we can rewrite (4.69) as follows

$$A_{pq} = c\left\{\int_{\xi_1}^{2\pi}\int_{\eta_1}^{1} R_{pq}(\rho)e^{-jq\theta}\rho d\rho d\theta - \int_{\xi_2}^{2\pi}\int_{\eta_2}^{1} R_{pq}(\rho)e^{-jq\theta}\rho d\rho d\theta\right\}, \tag{4.70}$$

where $c = 2\overline{f} + 3\,\mathrm{TV}(f)$ and $0 \le \xi_1, \xi_2 \le 2\pi$, $0 \le \eta_1, \eta_2 \le 1$. Assuming, without loss of generality, that $\xi_1 \le \xi_2$ and $\eta_1 \le \eta_2$ we can conclude that (4.70) is equal to

$$c\left\{\int_{\xi_1}^{\xi_2}\int_{\eta_1}^{1} R_{pq}(\rho)e^{-jq\theta}\rho d\rho d\theta + \int_{\xi_2}^{2\pi}\int_{\eta_1}^{\eta_2} R_{pq}(\rho)e^{-jq\theta}\rho d\rho d\theta\right\}. \tag{4.71}$$

Noting that $\mathbf{1}_{[\xi_1,\xi_2]}(\theta)$ and $\mathbf{1}_{[\xi_1,2\pi]}(\theta)$ are functions of bounded variation on $[0,2\pi]$ and using the result in [9] we have

$$\left|\int_{\xi_1}^{\xi_2} e^{-jq\theta}d\theta\right| = O\left(\frac{1}{q+1}\right) \text{ and } \left|\int_{\xi_1}^{2\pi} e^{-jq\theta}d\theta\right| = O\left(\frac{1}{q+1}\right). \tag{4.72}$$

By this and (4.71) it suffices to evaluate the integral

$$\int_0^{\eta} R_{pq}(\rho)\rho d\rho, \tag{4.73}$$

for $0 \le \eta \le 1$.

Using (4.3) we find that (4.73) is equal to

$$\frac{1}{4}\int_{-1}^{2\eta^2-1}\left(\frac{t+1}{2}\right)^{|q|/2} P_s^{(0,|q|)}(t)dt,$$

where $s = (p-|q|)/2$. This in turn is equal to

$$0.5\eta^{|q|+2}\int_0^1 y^{|q|/2}P_s^{(0,|q|)}(2\eta^2 y - 1)dy.$$

Using the identity in [14, p. 848] and some algebra we can find that the last integral is of the order

$$O\left(\frac{1}{(p+1)\sqrt{p-|q|+1}}\right).$$

This combined with (4.72) yields

$$|A_{pq}| = O\left(\frac{1}{(q+1)(p+1)\sqrt{p-|q|+1}}\right).$$

Consequently

$$\begin{aligned} \mathrm{TR}(\widehat{f}_T) &= \sum_{p=T+1}^{\infty}\sum_{q=-p}^{p}\frac{p+1}{\pi}|A_{pq}|^2 \\ &\leq c_1\sum_{p=T+1}^{\infty}\sum_{q=0}^{p}\frac{1}{(p+1)(p-q+1)(q+1)^2}. \end{aligned} \tag{4.74}$$

To evaluate the above double sum let us split the inner sum in (4.74) as follows

$$\sum_{q=0}^{p}\frac{1}{(p-q+1)(q+1)^2} = \sum_{q=0}^{p/2}\frac{1}{(p-q+1)(q+1)^2} \tag{4.75}$$

$$+\sum_{q=p/2+1}^{p}\frac{1}{(p-q+1)(q+1)^2}. \tag{4.76}$$

It is clear that the first term on the right-hand side of (4.76) is bounded by

$$\frac{1}{(p/2+1)}\sum_{q=0}^{\infty}\frac{1}{(q+1)^2} = \frac{\pi^2}{3}\frac{1}{p+2}.$$

The second term in (4.76) is bounded by

$$\sum_{q=p/2+1}^{\infty}\frac{1}{(q+1)^2} \leq \frac{2}{p+2}.$$

Consequently the sum on the right-hand side of (4.74) does not exceed

$$\left(\frac{\pi^2}{3}+2\right)\sum_{p=T+1}^{\infty}\frac{1}{(p+1)(p+2)} \leq \left(\frac{\pi^2}{3}+2\right)\frac{1}{T+1}.$$

This concludes the proof of Lemma 4.3. □

Chapter 5

Accurate Computation of Orthogonal Radial Moments

An effective numerical algorithm for high precision computation of the radial Zernike moments is presented. Thus far we have shown that there inevitably exist two (provided that noisy-free image is observed) kinds of errors limiting the precision of computation of radial moments, i.e., geometric error and numerical integration error. While the latter error is common to all moment based image analysis methods utilizing a certain type of discretization algorithms of an analog image, the former is unique to non-rectangular image domains. It is proved in this chapter that both geometric error and numerical integration error can be greatly reduced by adopting polar coordinates. This reduction is obtained at the expense of introducing the Cartesian-polar interpolation procedure which, however, has a negligible effect on the overall accuracy of the proposed method. We describe in detail our approach including the derivation of formulas for computation of Zernike moments under the polar coordinate system, the polar pixel arrangement scheme and the image conversion via interpolation. The effect of image interpolation on the accuracy of Zernike moments is also analyzed. Both theoretical and numerical comparisons are made between the proposed method and commonly used techniques. The results show the polar approach to be greatly advantageous.

5.1. Introduction

As we have noted in the previous chapters the radial Zernike type orthogonal moments (ZMs) possess distinguished characteristics such as the magnitude invariance to image rotations and reflections making them an excellent choice for a

number of applications in pattern recognition, image analysis, digital watermarking, medical imaging, and ophthalmology. In all these applications, inaccurate computation methods for ZMs derivation have been employed. Some fundamental aspects of the Zernike moments have been investigated in Chapters 3 and 4 concerning their accuracy and the reconstruction power. There, it has been demonstrated that for noisy free images we have two types of inherent errors, namely geometric error and numerical integration error. It is clear that these errors limit some attractive properties of ZMs. For instance, the property of magnitude invariance to image rotation depends heavily on the accuracy of ZMs. With the existing methods of computing ZMs, some of the moment magnitudes are not truly rotationally invariant.

In this chapter we introduce a highly accurate method for the Zernike moment computing based on the polar coordinates. A detailed numerical algorithm that is both computationally efficient and accurate is developed. It is shown that if ZMs are calculated in polar coordinates rather than in conventional Cartesian coordinates, the loss of accuracy due to geometric error and numerical integration error can be greatly reduced. We present a detailed description of the new approach including the derivation of formulas for ZMs computation in the polar coordinates system, the polar pixel arrangement scheme and the image conversion via interpolation. In our approach the error introduced due to the Cartesian-polar image interpolation has a minimal effect on the overall accuracy of the proposed method. As a result, the invariance properties of ZMs are greatly preserved. Furthermore, we can use, without any loss in the accuracy, ZMs of very high orders which may have important consequences in patter recognition problems.

This chapter is organized as follows. In Section 5.2, we discuss the accuracy issue in terms of geometric and numerical integration errors in the conventional method of ZMs computation. In Section 5.3, the main idea of the polar approach is formulated. This includes the principles of the algorithm, a novel and practical disk partitioning scheme to facilitate our method, and a proper image resampling technique. Theoretical and empirical evaluation of our algorithm is given in Section 5.4 and Section 5.5, respectively. The summary and concluding remarks are given in Section 5.6.

It is also worth noting that in the literature dealing with pattern recognition strong emphasis has been given to the problem of finding fast algorithms for calculation of moment descriptors [114,159,160]. We believe that the accuracy issue should be addressed before introducing any fast algorithm.

5.2. Traditional Methods of Zernike Moment Computation

Let us recall that for a continuous image function $f(x, y)$, the Zernike moment of order n with repetition m is defined as

$$A_{nm} = \iint_D f(x, y) V^*_{nm}(x, y) dxdy, \tag{5.1}$$

where the double integral is performed over the unit disk $D = \{(x, y) : x^2 + y^2 \leq 1\}$, and $V_{nm}(x, y)$ is a Zernike function defined as

$$V_{nm}(x, y) = R_{nm}(\rho) e^{jm\theta}, \tag{5.2}$$

where $\rho = \sqrt{x^2 + y^2}$, $\theta = \tan^{-1}(y/x)$. Here n is a non-negative integer and m is an integer such that $n - |m|$ is even and non-negative. The radial Zernike polynomial $R_{nm}(\rho)$ is defined by the following formula

$$R_{nm}(\rho) = \sum_{s=0}^{(n-|m|)/2} \frac{(-1)^s (n-s)! \rho^{n-2s}}{s! \left(\frac{n+|m|}{2} - s\right)! \left(\frac{n-|m|}{2} - s\right)!}. \tag{5.3}$$

For digital images, (5.1) cannot be applied directly. Now assume that an $N \times N$ image matrix $F(i, j), i = 1, ..., N, j = 1, ..., N$, is given. In order to compute its Zernike moments, one has to map $F(i, j)$ onto another function $f(x_i, y_j)$ defined on $[-1, 1]^2$, such that $f(x_i, y_j) = F(i, j), i = 1, ..., N, j = 1, ..., N$, where $x_i = (2i - N - 1)/N$ and $y_j = (2j - N - 1)/N$. Without loss of generality, it is assumed that $f(x_i, y_j)$ is a function with all its pixels inside the unit circle. The following formula for approximating A_{nm} has been proposed in Chapter 3

$$\widetilde{A}_{nm} = \sum_i \sum_j w_{nm}(x_i, y_j) f(x_i, y_j), \tag{5.4}$$

where the values of i and j are taken such that $(x_i, y_j) \in D$, as shown in Fig. 5.1, and

$$w_{nm}(x_i, y_j) = \int_{x_i - \frac{\Delta}{2}}^{x_i + \frac{\Delta}{2}} \int_{y_j - \frac{\Delta}{2}}^{y_j + \frac{\Delta}{2}} V^*_{nm}(x, y) dxdy, \tag{5.5}$$

where $\Delta = 2/N$ is the pixel width/height. For the computation of the factor $w_{nm}(x_i, y_j)$, some methods of numerical integration can be applied. The most commonly used formula, which is apparently the simplest and least accurate, is the following

$$w_{nm}(x_i, y_j) \approx \Delta^2 V^*_{nm}(x_i, y_j). \tag{5.6}$$

Some more complex numerical techniques, based on the multidimensional cubature algorithms, were discussed in Chapter 3 in order to improve the accuracy of computing $w_{nm}(x_i, y_j)$.

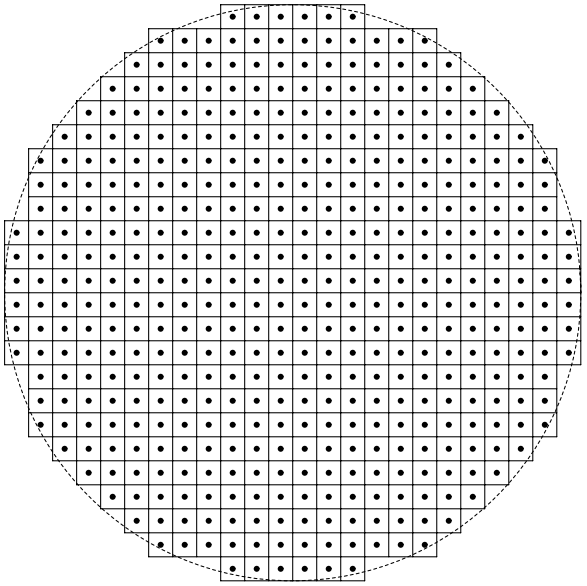


Figure 5.1: An illustration of the Cartesian pixel grid for computing Zernike moments

Nevertheless, as pointed out in Chapter 3, the accuracy of Zernike moments computed by formula (5.4) suffers from two sources of errors, namely geometric error and numerical integration error. The former is due to the fact that the total area covered by all the square pixels involved in the computation of Zernike moments via (5.4) is not exactly the unit disk, as illustrated by the ragged border in Fig. 5.1. The latter results from the numerical integration by an approximation formula like (5.6). Although some techniques can be deployed (see Chapter 3) to alleviate the inherent accuracy problem, the aforementioned errors can never be eradicated provided that the computation of Zernike moments is performed with the Cartesian coordinate system.

5.3. Zernike Moments in Polar Coordinates

The cause of errors in computing ZMs by formula (5.4) lies in the adoption of Cartesian coordinates for the computation, which is justified by the fact that digital images are represented by square pixels. However, this practice of computing

does not take into account the circular nature of Zernike polynomials. In this section we present an algorithm for computation of ZMs in polar coordinates, in which neither geometric error nor numerical integration error (represented by the integral in (5.5)) is present.

Principles

To remove geometric error and numerical integration error in ZMs computation, we need to take into account the circular nature of the Zernike functions. It is intuitive that geometric error can be avoided by using non-square pixels, whose areas add up to that of the unit disk. Furthermore, we can use an analytical method instead of numerical approximations for the pixel-wise integration of basis functions, which is accurate and efficient in moment computation [45]. Equation (5.2) shows that Zernike polynomials can be naturally expressed in polar coordinates (ρ, θ). This suggests that adoption of polar coordinates could facilitate the computation of ZMs. For this purpose, we rewrite the definition of Zernike moment (5.1) in its equivalent form based on polar coordinates

$$A_{nm} = \int_0^{2\pi} \int_0^1 \widetilde{f}(\rho, \theta) R_{nm}(\rho) e^{-jm\theta} \rho d\rho d\theta, \tag{5.7}$$

where $\widetilde{f}(\rho, \theta) = f(\rho \cos\theta, \rho \sin\theta)$. Let the image $\widetilde{f}(\rho, \theta)$ be approximated by a piecewise constant function $\widehat{f}(\rho, \theta)$ defined over a set of concentric sectors Ω_{uv} which meet the following two criteria

$$\bigcup_{(u,v)} \Omega_{uv} = D, \tag{5.8}$$

and

$$\Omega_{uv} \cap \Omega_{u'v'} = \emptyset \quad \forall\, (u, v) \neq (u', v'). \tag{5.9}$$

Consequently we can define the following approximate of A_{nm}

$$\widehat{A}_{nm} = \sum_u \sum_v \widehat{f}(\rho_{uv}, \theta_{uv}) \omega_{nm}(\rho_{uv}, \theta_{uv}), \tag{5.10}$$

where $\widehat{f}(\rho_{uv}, \theta_{uv})$ is the estimated image intensity over Ω_{uv}, centered at (ρ_{uv}, θ_{uv}), and the double summation is performed over all the sectors inside the unit disk.

The factor $\omega_{nm}(\rho_{uv}, \theta_{uv})$ is an integral over Ω_{uv}, i.e.,

$$\begin{aligned}\omega_{nm}(\rho_{uv}, \theta_{uv}) &= \iint_{\Omega_{uv}} R_{nm}(\rho) e^{-jm\theta} \rho d\rho d\theta \\ &= \int_{\rho_{uv}^{(s)}}^{\rho_{uv}^{(e)}} R_{nm}(\rho) \rho d\rho \int_{\theta_{uv}^{(s)}}^{\theta_{uv}^{(e)}} e^{-jm\theta} d\theta, \end{aligned} \tag{5.11}$$

where $\rho_{uv}^{(s)}$ and $\rho_{uv}^{(e)}$ denote the starting and ending radii of Ω_{uv} respectively, while $\theta_{uv}^{(s)}$ and $\theta_{uv}^{(e)}$ denote the starting and ending angles of Ω_{uv} respectively. Fig. 5.2 illustrates the variables introduced above. The formula in (5.11) is the product of two integrals, whose exact values can be easily obtained analytically.

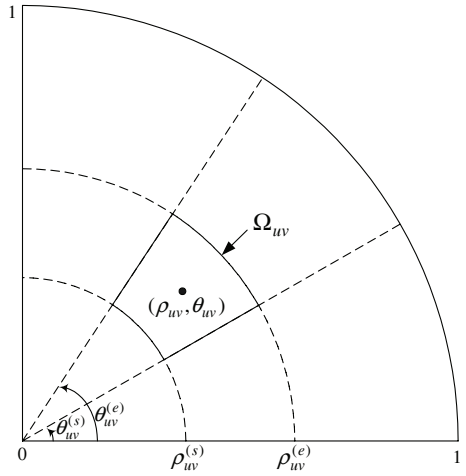


Figure 5.2: The sector Ω_{uv} representing a polar pixel. The point (ρ_{uv}, θ_{uv}) defines the location of Ω_{uv}, where $\rho_{uv} = \left(\rho_{uv}^{(s)} + \rho_{uv}^{(e)}\right)/2$ and $\theta_{uv} = \left(\theta_{uv}^{(s)} + \theta_{uv}^{(e)}\right)/2$

Hence by virtue of (5.3) we have

$$\begin{aligned}&\int_{\rho_{uv}^{(s)}}^{\rho_{uv}^{(e)}} R_{nm}(\rho) \rho d\rho \\ &\qquad = \sum_{\ell=0}^{(n-|m|)/2} \frac{(-1)^\ell (n-\ell)! [(\rho_{uv}^{(e)})^{n-2\ell+2} - (\rho_{uv}^{(s)})^{n-2\ell+2}]}{(n-2\ell+2)\ell! (\frac{n+|m|}{2} - \ell)! (\frac{n-|m|}{2} - \ell)!}. \end{aligned} \tag{5.12}$$

and

$$\int_{\theta_{uv}^{(s)}}^{\theta_{uv}^{(e)}} e^{-jm\theta} d\theta = \begin{cases} \frac{j}{m} [e^{-jm\theta_{uv}^{(e)}} - e^{-jm\theta_{uv}^{(s)}}], & m \neq 0, \\ \theta_{uv}^{(e)} - \theta_{uv}^{(s)}, & m = 0. \end{cases} \tag{5.13}$$

Combining (5.10), (5.11), (5.12) and (5.13), we obtain the estimated formula for ZMs of $f(\cdot,\cdot)$, without introducing any geometric and numerical integration errors.

A Polar Pixel Structure for ZMs Computation

As we have already noted, in order to eliminate geometric and numerical integration errors, we must represent $\{f(x_i, y_j)\}$ over the polar sectors satisfying conditions (5.8) and (5.9). If we imagine each sector as a fan-shaped pixel whose value is determined by that of its central point, then the question arises of what should be the arrangement of these fan-shaped pixels. There are numerous schemes satisfying conditions (5.8) and (5.9). An example of the most obvious structure is shown in Fig. 5.3, in which the unit disk is uniformly divided along both the radial and angular directions. This scheme has the advantage of easy implementation. However, it behaves poorly in terms of image representation. This is due to the fact that the areas of the sectors vary greatly, and it is impossible to achieve both efficiency and accuracy of information representation. In fact, if the inner sectors are required to be large enough to represent image information efficiently, the outer sectors are too large to accurately represent the original image information. On the other hand, if the outer sectors are required to be small enough to accurately represent the image information, then there are too many inner sectors for the scheme to be efficient.

To overcome the aforementioned problems of the partition scheme shown in Fig. 5.3, we need to design a more appropriate tiling. The following criteria are used for finding a suitable polar pixel partition of the image plane.

- All the sector areas should be approximately equal. Note that in Cartesian coordinates all pixels are of equal size.
- The number of polar pixels inside the unit circle should not be smaller than that of the Cartesian pixels inside the unit circle, so that the necessary image resolution could be maintained without loss of information.
- The polar pixels should be as "squared" as possible, i.e., the lengths of the borders of a sector should be close enough, so that the image distortion due to the coordinate system conversion could be kept at a low level.
- In order to facilitate the storage and computation processes, the polar pixel structure should be as simple and regular as possible,

Following these guidelines, we propose a pixel arrangement scheme illustrated in Fig. 5.4. The details of this structure are listed below.

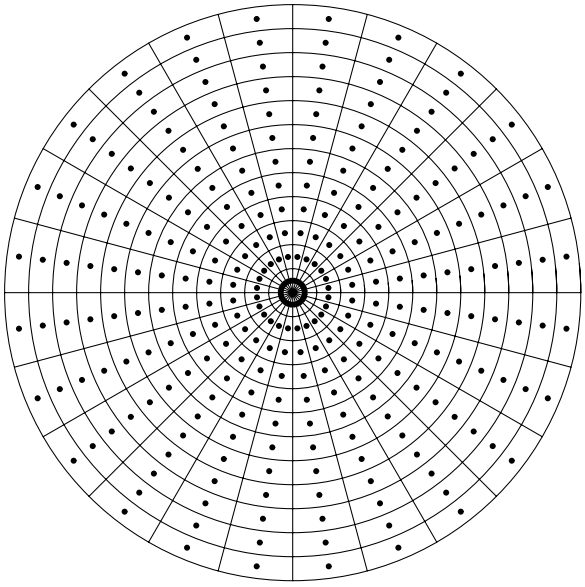


Figure 5.3: An example of a tentative polar pixel grid for efficient computation of Zernike moments

- The unit disk is uniformly divided along the radial direction into U sections, with the separating circles located at $\{k/U, k = 1, ..., U\}$.
- The k-th ring-shaped section is equally divided into $(2k-1)V$ sectors by radii starting from the origin, with angles $\{(i-1)2\pi/((2k-1)V), i = 1, ..., (2k-1)V\}$. V is the number of sectors contained in the innermost section.

It can be shown by a simple algebra that the unit disk is divided into VU^2 sectors, each of which has an area of $\pi/(VU^2)$. The values of U and V should be set properly. A low value of VU^2 is advantageous in terms of computation and implementation, but may represent inadequately the image information. On the other hand, a high value of VU^2 is beneficial for representation of the image, but entails heavy workload. In practice, we recommend setting $V = 4$ and $N/2 \leq U \leq N$ for an $N \times N$ image.

Equipped with the above introduced scheme of the polar pixel arrangement, the formula for ZM computation in (5.10) can be further rewritten as

$$\widehat{A}_{nm} = \sum_{u=1}^{U} \sum_{v=1}^{(2u-1)V} \widehat{f}(\rho_{uv}, \theta_{uv}) \iint_{\Omega_{uv}} R_{nm}(\rho) e^{-jm\theta} \rho d\rho d\theta, \tag{5.14}$$

where the integral can be explicitly evaluated using (5.12) and (5.13).

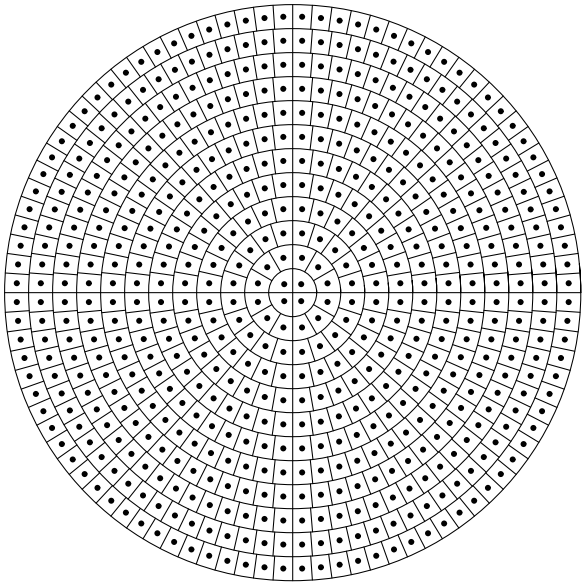


Figure 5.4: The proposed structure of polar pixel tiling for efficient computation of Zernike moments

Image Representation in Polar Coordinates

We have seen in the previous section that computing ZMs in polar coordinates results in neither geometric error nor numerical integration error, provided that the image plane is partitioned by a certain polar pixel structure as in Fig. 5.4. However, in practice, a digital image is defined by a set of square Cartesian pixels, as shown in Fig. 5.1. It can be verified that the locations of most of the polar pixels do not coincide with those of the Cartesian pixels. Therefore, we have to derive the polar counterpart of a given Cartesian image before computing its ZMs in polar coordinates. This issue can be resolved by applying an image interpolation procedure.

There are a number of existing image interpolation techniques [16, 120, 153] which we can use to determine the values of polar pixels. The simplest and least accurate one is the nearest point interpolation, by which the value of a polar pixel is set to that of the closest Cartesian pixel. Another technique is the bilinear interpolation, which determines the value of a point by the linearly weighted average of the four neighboring pixel values. The bilinear technique yields better image quality, and has been widely used in many applications. A more advanced approach, which we adopt in our work, is bicubic interpolation

[21, 120, 153]. In this method the 16 neighboring pixels are used to compute the value of an interpolated point. This usually yields a smooth image with very small interpolation error. The 1-D kernel function representing the bicubic interpolation is a cubic spline

$$h(x) = \begin{cases} \frac{3}{2}|x|^3 - \frac{5}{2}|x|^2 + 1, & |x| \leq 1, \\ -\frac{1}{2}|x|^3 + \frac{5}{2}|x|^2 - 4|x| + 2, & 1 < |x| \leq 2, \\ 0, & \text{otherwise.} \end{cases} \tag{5.15}$$

The interpolated value of $\{f(x_i, y_j)\}$ at Ω_{uv} can be obtained based on the 2-D cubic convolution between the image function $\{f(x_i, y_j)\}$ and the kernel $h(x)h(y)$, i.e.,

$$\widehat{f}(\rho_{uv}, \theta_{uv}) = \sum_{i=a-1}^{a+2} \sum_{j=b-1}^{b+2} f(x_i, y_j) h\left(\frac{\rho_{uv}\cos\theta_{uv} - x_i}{\Delta}\right) h\left(\frac{\rho_{uv}\sin\theta_{uv} - y_j}{\Delta}\right), \tag{5.16}$$

where $a = \left\lfloor \frac{\rho_{uv}\cos\theta_{uv}}{\Delta} \right\rfloor + N/2$ and $b = \left\lfloor \frac{\rho_{uv}\sin\theta_{uv}}{\Delta} \right\rfloor + N/2$, and $\Delta = 2/N$ is the pixel width.

Clearly we could apply more advanced interpolation methods which require an additional pre-filtering step [16]. This might further improve the accuracy of our method at the expense of increased computational workload. Simulation results presented in this chapter indicate that the bicubic interpolant gives a proper trade-off between the accuracy and the computational complexity.

5.4. The Accuracy Analysis of the Algorithm

As we have shown in the previous section, the proposed polar approach avoids entirely both geometric error and numerical integration error in computing Zernike moments. If a digital image is represented directly by a grid of pixels structured as in Fig. 5.4, its Zernike moments can be computed accurately by the proposed algorithm without incurring on the aforementioned errors. In practice, however, an image is defined over the square grid and it is necessary to convert it into its polar coordinates counterpart. This conversion is done through an interpolation procedure like the formula in (5.16) resulting in some interpolation error.

In this section we describe the influence of the interpolation error on the accuracy of the estimate $\widehat{A}_{nm}$ defined in (5.10). To this end let us consider the

estimate $\widehat{A}_{nm}$ with $\widehat{f}(\rho,\theta)$ being any interpolation scheme which has the following degree of accuracy

$$\max_{(\rho,\theta)} |\widehat{f}(\rho,\theta) - \widetilde{f}(\rho,\theta)| \leq C\Delta^p, \tag{5.17}$$

where Δ is the pixel width and $p \geq 1$ characterizes the degree of accuracy of the interpolant. In (5.17) the constant C depends on the smoothness of the image function $f(\cdot,\cdot)$ and an interpolation kernel of the applied interpolant. It is known that $p = 1$ corresponds to the nearest-neighbour algorithm, and $p = 2$ to the linear interpolation, whereas the cubic convolution method meets (5.17) with $p = 3$, [21, 120, 153]. It is worth mentioning that pre-filtering interpolation algorithms [16] can satisfy (5.17) with $p > 3$, provided that appropriate smoothing conditions of the image function are met.

Let us consider the estimate $\widehat{A}_{nm}$ in (5.10) with the general linear interpolation scheme $\widehat{f}(\rho_{uv},\theta_{uv})$. We wish to evaluate the size of the difference between $\widehat{A}_{nm}$ and the true Zernike moment A_{nm}. To do so we assume that the analog image function $f(x,y)$ is in the class of functions of bounded variation on D. The following theorem shows that the difference $\widehat{A}_{nm} - A_{nm}$ is of order $\max(\Delta^p, \Delta)$.

Theorem 5.1. *Let $f(\cdot,\cdot)$ be a function of bounded variation on D. Let $\widehat{A}_{nm}$, defined in (5.10), be the estimate of A_{nm} utilizing the general interpolation scheme which satisfies assumption (5.17). Then we have*

$$|\widehat{A}_{nm} - A_{nm}| \leq \left(\frac{\pi}{n+1}\right)^{1/2} \left\{8C^2\Delta^{2p} + 4f_{\max}\,\mathrm{TV}(f)\Delta^2\right\}^{1/2}, \tag{5.18}$$

where $\mathrm{TV}(f)$ is the total variation of $f(\cdot,\cdot)$ over D and $f_{\max} = \max_{(x,y)\in D} f(x,y)$.

The proof of this theorem can be found in the Appendix. The bound in (5.18) contains two unrelated terms. The first one, being of order $O(\Delta^p)$, characterizes the applied interpolation scheme of order p. On the other hand, the second term in (5.18) is of order $O(\Delta)$ and describes the discretization error in replacing the integral in (5.7) by the sum appearing in (5.10). This term can be reduced by putting some further smoothness conditions on $f(\cdot,\cdot)$. In fact, if $f(\cdot,\cdot)$ is differentiable then the term $O(\Delta^2)$ in (5.18) is replaced by the term of order $O(\Delta^3)$. Hence under the assumption of Theorem 5.1 we have

$$\widehat{A}_{nm} = A_{nm} + C_1(n+1)^{-1/2}\Delta, \tag{5.19}$$

for some positive constant C_1 depending on $f(\cdot,\cdot)$.
Thus we can conclude that the interpolation error is of the second order of importance to the discretization error.

It is worth noting that due to the result in Chapter 3, the square grid based method given in (5.4) and (5.5) exhibits the following error

$$\widetilde{A}_{nm} = A_{nm} + C_2(n+1)^{-1/2}\Delta + C_3 \min((n+1)^{-1/2}\Delta^{\alpha/2}, \Delta^{\alpha}), \tag{5.20}$$

where $1 \leq \alpha < 3/2$ is the exponent characterizing the lattice points approximation of a circle, with the best known $\alpha = 285/208$, see Chapter 3. C_2, C_3 are positive constants depending on $f(\cdot,\cdot)$. For image functions possessing the first derivative, the result in (5.19) improves to $O(\Delta^{3/2})$, whereas the second term on the right-hand-side of (5.20) remains the same.

Thus, we have obtained the qualitative result that the proposed method is more accurate than any method based on the square partition of the image plane.

5.5. Empirical Evaluation of the Algorithm Accuracy

In this section, the empirical effectiveness of the proposed polar approach is investigated. We illustrate it from three different perspectives, namely, the moment magnitudes of a constant image, the image reconstruction from a finite set of computed ZMs, and ZM magnitude invariance to image rotation.

Improvement of Zernike Moment Accuracy

To illustrate the accuracy of ZM computation, we use a 128×128 image with a constant intensity value 127 as the test image, to which both the conventional Cartesian method defined by (5.4) and (5.6) and the proposed polar method are applied. The reason we choose the constant image is that, in theory, all its true ZMs $A_{nm} = 0$ except that $A_{0,0} = 127$. Therefore by looking at the magnitudes of the computed moments, we can assess the accuracy of the algorithm. Considering ZMs' symmetry property, $A_{nm} = A^*_{n,-m}$, here we are only concerned with the A_{nm} with $m \geq 0$.

First, the constant image's ZMs with orders between 1 and 40 are calculated based on the conventional Cartesian method, and placed in the following order

$$\{\widetilde{A}_{1,1}, \widetilde{A}_{2,0}, \widetilde{A}_{2,2}, \widetilde{A}_{3,1}, \widetilde{A}_{3,3}, \widetilde{A}_{4,0}, \widetilde{A}_{4,2}, \ldots, \widetilde{A}_{40,40}\}.$$

Fig. 5.5 depicts the magnitudes of these ZMs as a function of their index numbers. It can clearly be seen that many ZMs computed in this way have considerable magnitudes, which reflect the inaccuracy of the Cartesian method.

Then, the same group of ZMs is calculated with the proposed polar method. The result is shown in Fig. 5.6, in which the magnitudes of all the moments are below $O(10^{-10})$. Comparing Fig. 5.5 and Fig. 5.6, one immediately concludes that the proposed approach is greatly superior to the conventional method.

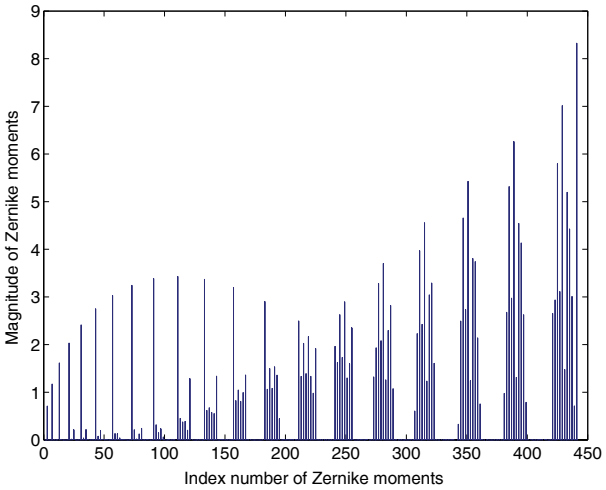


Figure 5.5: The magnitudes of ZMs with orders between 1 and 40, of a 128×128 constant image function, computed with the Cartesian method

Improvement of Image Reconstruction

Image reconstruction from a finite number of moments, L, can be performed with the following formula:

$$\widehat{f}(x,y) = \sum_{i=1}^{L} [\widehat{A}_{n_i m_i} V_{n_i m_i}(x,y) + \mathbf{1}(m_i \neq 0) \widehat{A}_{n_i,-m_i} V_{n_i,-m_i}(x,y)], \qquad (5.21)$$

where $\mathbf{1}(\cdot)$ is the indicator function. It was shown in Chapter 4 that the reconstruction error consists of two parts. One part of the error is due to the finite value of L, and the other comes from the inaccuracy of the computed ZMs $\widehat{A}_{nm}$. The former can be reduced by increasing L, while the latter was shown to be inevitable due to the inherent geometric and numerical integration errors [145, 180]. As we have already shown above, the loss of moment accuracy brought by geometric error and numerical integration error that used to be inherent in the Cartesian method can now be significantly reduced in polar coordinates. Therefore the reconstruction error due to the inaccuracy of $\widehat{A}_{nm}$ can be greatly improved. This has been verified by simulation results. We use the 128×128 Lena image, shown in Fig. 5.7a, to illustrate the image reconstruction performance. The Zernike moments up to the order of 200 are computed with the conventional Cartesian approach and the proposed polar approach, respectively. Then the computed moments are used to reconstruct images. Fig. 5.8 depicts some of the images reconstructed with the Cartesian system-based moments, while Fig. 5.9 displays

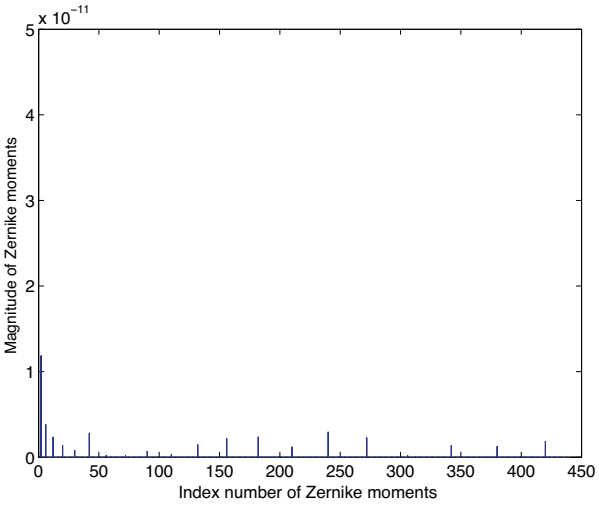


Figure 5.6: The magnitudes of ZMs with orders between 1 and 40, of a 128 × 128 constant image function, computed with the proposed polar method

(a) **(b)**

Figure 5.7: The test images. (a) Original 128 × 128 image of Lena. (b) Lena image rotated through 15°

the images reconstructed with the polar system-based moments. It can be seen from Fig. 5.8 that some erroneous pixels along the border of the unit circle are very obtrusive in the images based on Cartesian ZMs, and the number of erroneous pixels increases quickly as the moment order increases. However, such erroneous pixels do not exist in the images based on Polar ZMs.

To compare qualitatively the performances of the two approaches in terms of image reconstruction, we have experimented on more different numbers of ZMs for image recovery. Specifically, ZMs up to orders $\{n = 2i\}_{i=0}^{100}$ were used to reconstruct the image respectively. The quality of each reconstructed image is

Figure 5.8: Image reconstruction from ZMs computed in the Cartesian coordinate system. First row from left to right: images reconstructed from ZMs up to order 20, 40, 60, 80 and 100 respectively. Second row from left to right: images reconstructed from ZMs up to order 120, 140, 160, 180 and 200, respectively

measured in terms of peek signal-to-noise ratio (PSNR), which is defined by

$$\mathrm{PSNR}(f, \widehat{f}) = 10 \log_{10} \frac{f_{\max}^{\,2}}{\sigma_e^2}, \tag{5.22}$$

where f is the original image and $\widehat{f}$ is the reconstructed image, both with dimensions $N \times N$, $f_{\max}$ is the maximum pixel value of image f, and

$$\sigma_e^2 = \frac{1}{N^2} \sum_{i=1}^{N} \sum_{j=1}^{N} [\widehat{f}(x_i, y_j) - f(x_i, y_j)]^2 \tag{5.23}$$

is the average squared error. The test results are shown in Fig. 5.10, from which two important conclusions can be drawn. Firstly, for low orders of ZMs, approximately $n < 20$, the quality of the reconstructed images via the polar ZMs is similar to that of the reconstructed images via the Cartesian ZMs. But as n becomes larger, the former gets significantly better than the latter. Secondly, the quality of the polar ZM-reconstructed images increases monotonically with n. However, in the Cartesian case, as n increases to a certain point, approximately 40, the image quality reaches its maximum value. After this maximum value, the image quality deteriorates. This is because the reconstruction error incurred by geometric and numerical integration errors increases with n, and at some point it outweighs the quality gain from the population increase of ZMs [145, 180]. Such picking effect is also present for the proposed method but the deterioration of the image quality appears for a very high moment order.

Figure 5.9: Image reconstruction from ZMs computed in the polar coordinate system. First row from left to right: images reconstructed from ZMs up to order 20, 40, 60, 80 and 100 respectively. Second row from left to right: images reconstructed from ZMs up to order 120, 140, 160, 180 and 200, respectively

Improvement of ZMs Rotational Invariance

If an image $f(x, y)$ is rotated through α degrees counterclockwise we have demonstrated in Chapter 3 that the Zernike moment of the resulting image, $A_{nm}^{(\alpha)}$, is related to that of the original image by

$$A_{nm}^{(\alpha)} = A_{nm} e^{-jm\alpha}. \tag{5.24}$$

This leads to the well known rotational invariance property $|A_{nm}^{(\alpha)}| = |A_{nm}|$, which holds if ZMs are accurate. However, if ZMs are computed from discrete data, this property holds only approximately. Therefore, an inspection of ZM's magnitude change before and after image rotation reveals the accuracy of moment computation. Let $|\widehat{A}|$ and $|\widehat{A}^{(r)}|$ denote the estimated moment magnitudes of the digital image before and after rotation, respectively. We are interested in the error $\Delta\widehat{A} = |\widehat{A}^{(r)}| - |\widehat{A}|$. As an example, the image of Lena is rotated through 15° to become the image shown in Fig. 5.7b. Fig. 5.11a depicts $\Delta\widetilde{A}$ in the case of the Cartesian approach, while Fig. 5.11b illustrates $\Delta\widehat{A}$ resulting from the polar approach. It is clear that the polar ZMs greatly outperform the Cartesian ZMs in terms of the rotational invariance property.

To quantitatively evaluate the improvement of magnitude invariance, we define the mean-square-error of ZMs magnitudes as

$$\text{MSE}(|\widehat{A}|, |\widehat{A}^{(r)}|) = \frac{1}{L} \sum_{i=1}^{L} (|\widehat{A}_{n_i m_i}| - |\widehat{A}_{n_i m_i}^{(r)}|)^2, \tag{5.25}$$

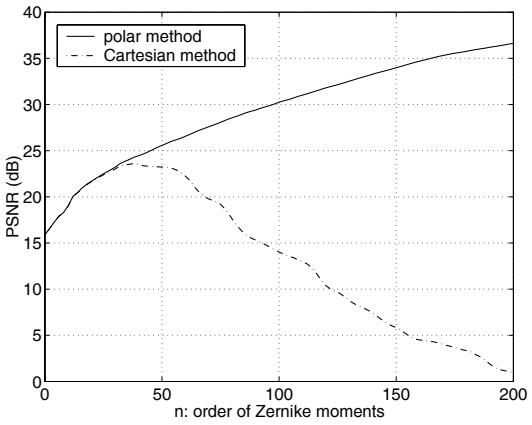


Figure 5.10: The image reconstruction quality in terms of PSNR as a function of the order of Zernike moments

where L is the number of ZMs involved in the evaluation. Depicted in Fig. 5.12 are $\mathrm{MSE}(|\widehat{A}|, |\widehat{A}^{(r)}|)$ in the cases of both the Cartesian approach and the polar approach. We have experimented with rotation angles from 0° to 90°, with a 2.25° interval. For each rotation angle, we computed the first 256 ZMs of the rotated image with both Cartesian approach and polar approach respectively, and then obtained the corresponding MSE values according to (5.25). The advantage of the polar approach is evident. To be more specific, the ratio of two average MSEs, $\overline{\mathrm{MSE}}_{\mathrm{Cartesian}}/\overline{\mathrm{MSE}}_{\mathrm{polar}} \approx 61.66$.

5.6. Conclusions

In this chapter we have proposed a novel approach for high precision computation of Zernike moments for digital images. In contrast to the traditional Cartesian coordinates-based methods, this approach is designed within the polar coordinate system. Detailed aspects of the algorithm, such as the lattice structure of polar pixels and the generation of the polar image from its Cartesian counterpart, have been investigated. It was shown that with this algorithm, the two inherent kinds of errors from which the Cartesian method suffers are greatly reduced. For a digital image given in a Cartesian format, the accuracy of its Zernike moments is determined by the interpolation scheme involved. Our empirical and theoretical results show that the interpolation error is of the second order importance in the overall error.

A series of experimental results were described, and they have verified the

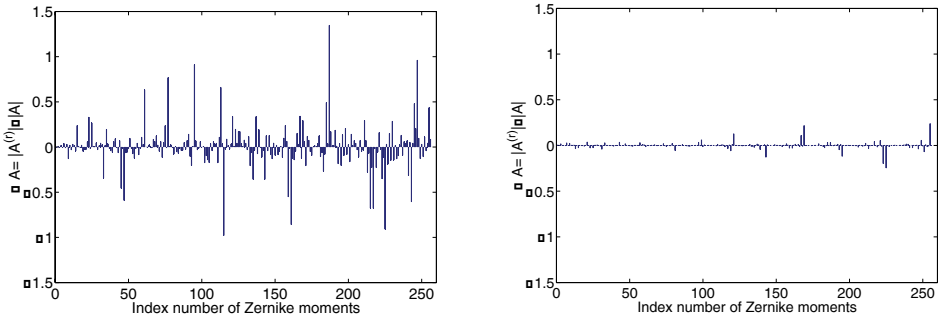


Figure 5.11: The magnitude differences between the first 256 ZMs of the two images shown in Fig. 5.7. (a) The results from the conventional Cartesian approach. (b) The results from the proposed polar approach

superiority of the proposed approach. It is worth noting, however, that the presented results have been obtained for noise-free images. If we observe an image in the presence of noise our method should be slightly modified. Indeed, the interpolation procedure cannot be longer used since noise interpolation is a highly undesirable practice. Instead one must apply some kind of smoothing procedure, e.g., cubic smoothing splines should replace the bicubic interpolation method.

5.7. Appendix

Proof of Theorem 5.1

Let us first observe that

$$\widehat{A}_{nm} - A_{nm} = \sum_{u=1}^{U} \sum_{v=1}^{(2u-1)V} \iint_{\Omega_{uv}} [\widehat{f}(\rho_{uv}, \theta_{uv}) - \widetilde{f}(\rho, \theta)] R_{nm}(\rho) e^{-jm\theta} \rho d\rho d\theta. \quad (5.26)$$

By Cauchy-Schwarz inequality we obtain

$$|\widehat{A}_{nm} - A_{nm}| \leq \left\{ \iint_D |R_{nm}(\rho) e^{-jm\theta}|^2 \rho d\rho d\theta \right\}^{1/2} \times \left\{ \sum_{u=1}^{U} \sum_{v=1}^{(2u-1)V} \iint_{\Omega_{uv}} |\widehat{f}(\rho_{uv}, \theta_{uv}) - \widetilde{f}(\rho, \theta)|^2 \rho d\rho d\theta \right\}^{1/2}. \quad (5.27)$$

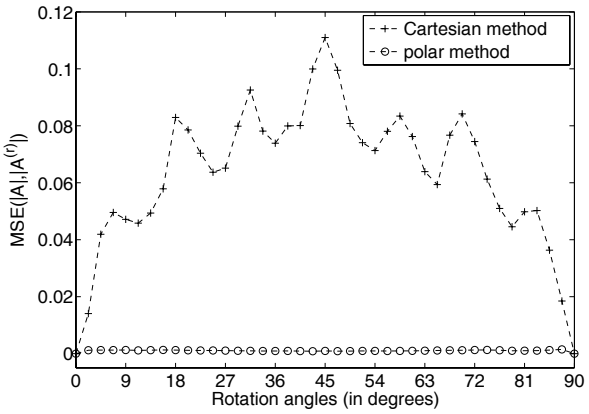


Figure 5.12: Mean-square-error of ZMs magnitudes of the images in Fig. 5.7a and Fig. 5.7b versus the rotation angle

Since $\iint_D |R_{nm}(\rho)e^{-jm\theta}|^2 \rho d\rho d\theta = \dfrac{\pi}{n+1}$, there remains the second term in (5.27) to be considered.

Let us notice that

$$\iint_{\Omega_{uv}} |\widehat{f}(\rho_{uv},\theta_{uv}) - \widetilde{f}(\rho,\theta)|^2 \rho d\rho d\theta \leq L_1 + L_2, \tag{5.28}$$

where

$$L_1 = 2\iint_{\Omega_{uv}} |\widehat{f}(\rho_{uv},\theta_{uv}) - \widetilde{f}(\rho_{uv},\theta_{uv})|^2 \rho d\rho d\theta, \tag{5.29}$$

and

$$L_2 = 2\iint_{\Omega_{uv}} |\widetilde{f}(\rho_{uv},\theta_{uv}) - \widetilde{f}(\rho,\theta)|^2 \rho d\rho d\theta. \tag{5.30}$$

By virtue of (5.17)

$$L_1 \leq 2C^2\Delta^{2p}\Delta^2, \tag{5.31}$$

where, without loss of generality, we assume that the area of Ω_{uv} is not greater than Δ^2.

Concerning the term L_2, we have

$$\begin{aligned} L_2 &\leq 2 \operatorname*{Osc}_{\Omega_{uv}}(\widetilde{f}) \iint_{\Omega_{uv}} |\widetilde{f}(\rho_{uv},\theta_{uv}) - \widetilde{f}(\rho,\theta)| \rho d\rho d\theta \\ &\leq 4 f_{\max} \operatorname*{Osc}_{\Omega_{uv}}(\widetilde{f}) \Delta^2, \end{aligned} \tag{5.32}$$

where

$$\underset{\Omega_{uv}}{\mathrm{Osc}}(\widetilde{f}) = \max_{(\rho,\theta),(\rho',\theta')\in\Omega_{uv}} \{|\widetilde{f}(\rho,\theta) - \widetilde{f}(\rho',\theta')|\}$$

is the oscillation of $\widetilde{f}(\rho,\theta)$ over the sector Ω_{uv}, and $f_{\max} = \max_{(x,y)\in D} f(x,y)$.

Substituting (5.28), (5.31) and (5.32) into (5.27), and noting that

$$\mathrm{TV}(f) = \sum_{u=1}^{U} \sum_{v=1}^{(2u-1)V} \underset{\Omega_{uv}}{\mathrm{Osc}}(\widetilde{f}) \tag{5.33}$$

we can obtain the following bound

$$|\widehat{A}_{nm} - A_{nm}| \leq \left(\frac{\pi}{n+1}\right)^{1/2} \left\{8C^2\Delta^{2p} + 4f_{\max}\,\mathrm{TV}(f)\Delta^2\right\}^{1/2}. \tag{5.34}$$

This proves the assertion of Theorem 5.1. □

Chapter 6

Testing Image Symmetry

In this chapter we examine an important application of the developed theory of radial moments, namely the problem of assessing whether an observed image may possess a certain type of symmetry. Formal statistical tests are constructed for finding a possible symmetry in an image observed under noise. We consider image invariance models with respect to reflections or rotations through rational angles, as well as joint invariance under both reflections and rotations. Furthermore, we propose a test for radiality, i.e., for checking whether the image function depends solely on the radius and not on the angle. These symmetry relations can be simply expressed as restrictions for the Zernike moments of the image function. Therefore, our test statistics are based on checking whether the estimated Zernike coefficients approximately satisfy the restrictions claimed. We derive the asymptotic distribution of the test statistics under both the hypothesis of symmetry and the fixed alternatives. The former result is used to construct asymptotic level-α tests for the hypothesis, while the latter can be employed for estimating the power of the test or for the purpose of validating symmetry. A theory of limiting distributions of generalized quadratic forms is employed.

6.1. Introduction

Symmetry plays an important role in image understanding and recognition. In fact, symmetric patterns are common in nature and the detection of an image symmetry can be utilized for designing effcient algorithms for object recognition, robotic manipulation, image animation, and image compression. It is clear that symmetric objects should be easier to recognize than non-symmetric ones and symmetry represents information redundancy which may be used to overcome occlusions and noise. For example, in the face recognition problem it would be

very useful to validate an approximate reflection symmetry with respect to a vertical axis of symmetry, Fig. 6.1 illustrates this concept.

Figure 6.1: Reflection symmetry of a human face

What does it mean to say that the image $f(x, y)$ is symmetric? Symmetry can be defined in terms of a combination of certain geometric transformations which when applied to a given planar image, yield a result indistinguishable from the original object. There are two basic symmetry types, namely rotational and reflection symmetries. Indeed, a finite non-periodic 2-D pattern may exhibit only these two kinds of symmetries.

An image shows rotational symmetry of order d, d being an integer, if it is invaraint to rotations of π/d, and its integer multiplies about the center of mass of the object. In particular, $d = 1$ corresponds to the requirement that

$$rf(x, y) = f(x, y), \tag{6.1}$$

where $rf(x, y) = f(-x, -y)$ is the image rotation through an angle π. Yet another important case of rotational symmetry is when $d = \infty$, i.e., when we have

$$\text{inv}\, f(x, y) = f(x, y), \tag{6.2}$$

where $\text{inv}\, f(x, y) = g(\sqrt{x^2 + y^2})$ for some univariate function $g(\cdot)$.

An image shows reflectional symmetry if it is invariant to reflection with respect to one or more lines called axes of symmetry. If there is only one axis of symmetry and it aligns with the y-axis the reflectional symmetry is defined as follows

$$\tau f(x, y) = f(x, y), \tag{6.3}$$

where $\tau f(x, y) = f(-x, y)$. Fig. 6.2 illustrates the aforementioned types of image symmetry.

(a) **(b)**

Figure 6.2: Examples of basic image symmetries: (a) reflection symmetry (b) rotation symmetry with $d = 3$.

A number of algorithms have been proposed for automatic detection and classification of rotational and reflectional symmetries in planar images [56, 125, 155, 238]. Most of the symmetry detectors utilize the image brightness or gradient information and they may be very sensitive to noise. Furthermore, the methods proposed assume that the image under examination is symmetric with respect to one of the aforementioned types. In practice, however, we do not have access to such a knowledge as we merely observe a noisy and digitized version of the original image. Thus, even if the original image is symmetric, its observed counterpart may not appear to be invariant to rotation and reflection operations. We therefore need formal statistical testing techniques to validate the existing symmetry in an image.

In this chapter we propose a systematic appraoch to testing an image symmetry utilizing the class of radial moments introduced in Chapter 3. In particular, we examine the symmetries defined in (6.1), (6.2), and (6.3). Furthermore testing the joint invariance under both reflections and rotations is investigated. Note that in the case of reflectional symmetry we make an assumption that the axis of symmetry (if an object is actually symmetric) is unique and defined by the y-axis. In general the axis of symmetry must be estimated and we refer to [56, 155] for efficient, although ad hoc, algorithms for finding all the axes of symmetry of symmetric planar objects.

We define our test statistics employing natural properties of the radial moments for the aforementioned symmetries. The statistics are in the form of the L_2 norm of the estimated radial moments with a suitable selected trunacation number. The established limit theorems allow us to design rigorous methods for testing symmetries of planar images observed in the presence of noise over a grid of pixels. Our tests are non-parametric as they do not need any prior knowledge

of the image shape and content. They are based on the region-based orthogonal moment descriptors and therefore in the case of rejecting the null hypothesis on the image symmetry they can still be used as a reconstruction method. This allows us to obtain a great deal of information about the image even if it has been classified to be non-symmetrical. This is a unique property of our testing method not shared by the existing algorithms for detecting symmetry in grey level images. Our theoretical developments leading to construction of formal statistical tests are based on the theory of the asymptotic behavior of quadratic forms of independent random variables.

The observation model of the image function is analogous to that given in Chapter 4. Hence let $f \in L_2(D)$ and take the noisy data

$$Z_{i,j} = f(x_i, y_j) + \epsilon_{i,j}, \quad (x_i, y_j) \in D, \quad 1 \le i, j \le n, \tag{6.4}$$

where the noise process $\{\epsilon_{i,j}\}$ is an i.i.d. random sequence with zero mean, finite variance $E\epsilon_{i,j}^2 = \sigma^2$ and finite fourth moment. We assume that the data are observed on a symmetric square grid of edge width Δ, i.e., $x_i - x_{i-1} = y_i - y_{i-1} = \Delta$ and $x_i = -x_{n-i+1}$, $y_i = -y_{n-i+1}$. Note that n is of order $1/\Delta$.

The question of symmetry of certain quantities is also relevant in statistical models. Such symmetry can simplify statistical inference, or it might be of interest in itself. For example, in linear as well as nonparametric regression models, it is important to know whether the error distribution is symmetrical around zero, since this can increase the efficiency of estimation as well as the quality of asymptotic approximations by the normal distribution, see, e.g., [34, 42, 95, 163]. Furthermore, one might also be interested in the symmetry of an unknown regression function, or in the symmetry of an unknown density of an i.i.d. sample of observations, and tests for these problems have been proposed in the literature as well [3, 138]. Nevertheless, testing symmetries in multi-dimensional functions has not been done yet.

6.2. The Zernike Orthogonal Basis

In this section we introduce relevant concepts and notation related to the Zernike polynomials and their approximation properties. We also define the corresponding image reconstruction formula. For extensive discussion of these concepts we refer to Chapters 2 and 4.

Zernike Polynomials and Function Approximation

The Zernike orthogonal polynomials are given by

$$V_{pq}(x,y) = R_{pq}(\rho)\, e^{jq\theta}, \quad (x,y) \in D, \tag{6.5}$$

where $\rho = \sqrt{x^2+y^2}$, $\theta = \arctan(y/x)$, j is the imaginary unit, and $R_{pq}(\rho)$ is the radial Zernike polynomial defined explicitly in Chapter 3. The indices (p,q) have to satisfy

$$p \geq 0, \quad |q| \leq p, \quad p - |q| \quad \text{even}, \tag{6.6}$$

and we will call such pairs (p,q) admissible. The Zernike polynomials satisfy the orthogonality relation, see Chapter 3, which implies that

$$\|V_{pq}\|^2 = \pi/(p+1) = n_p, \tag{6.7}$$

where $\|\cdot\|$ is the $L_2(D)$-norm.

In Chapter 3, see also [13], we noted that the Zernike polynomials are characterized as the unique orthogonal basis of $L_2(D)$ consisting of invariant polynomials of the general form (6.5), which contain a polynomial for every admissible pair (p,q) in (6.6), where p is the degree of $R_{pq}(\rho)$ and q is the index of angular dependence. Several useful properties of the radial Zernike polynomials $R_{pq}(\rho)$ have been listed in Chapter 3. In particular, we will need the property that $|R_{pq}(\rho)| \leq 1$.

Since the family $\{V_{pq}(x,y)\}$ for admissible (p,q) forms a complete and orthogonal system in $L_2(D)$, we can expand a function $f \in L_2(D)$ into a Fourier series

$$f(x,y) = \sum_{p=0}^{\infty} \sum_{q=-p}^{p} n_p^{-1}\, A_{pq}(f) V_{pq}(x,y). \tag{6.8}$$

Here and in the following the summation is only taken over admissible pairs (p,q). The norming factor n_p^{-1} arises due to (6.7), and the Fourier coefficient $A_{pq}(f)$ is defined by

$$A_{pq}(f) = \iint_D f(x,y)\, V_{pq}^*(x,y)\, dx\, dy.$$

Introducing the notation

$$\widetilde{f}(\rho,\theta) = f(\rho\cos\theta, \rho\sin\theta)$$

by using polar coordinates, one obtains

$$A_{pq}(f) = 2\pi \int_0^1 c_q(\rho, f)\, R_{pq}(\rho)\, \rho\, d\rho, \tag{6.9}$$

where

$$c_q(\rho, f) = \frac{1}{2\pi} \int_0^{2\pi} \widetilde{f}(\rho, \theta) e^{-iq\theta}\, d\theta.$$

We will also need Parseval's formula for the Zernike basis, i.e., for $f \in L_2(D)$ we have that

$$\|f\|^2 = \sum_{p=0}^{\infty} \sum_{q=-p}^{p} n_p^{-1} |A_{pq}(f)|^2. \tag{6.10}$$

Image Reconstruction

The observational model we employ in this chapter is similar to the one used in Chapter 4. Hence, let $f \in L_2(D)$ and suppose that we have data

$$Z_{i,j} = f(x_i, y_j) + \epsilon_{i,j}, \quad (x_i, y_j) \in D, \quad 1 \le i, j \le n, \tag{6.11}$$

where the noise process $\{\epsilon_{i,j}\}$ is an i.i.d. random sequence with zero mean, finite variance $E\epsilon_{i,j}^2 = \sigma^2$ and finite fourth moment, see (6.4).

The design points are illustrated in Fig. 3.4. As we have already noted, along the boundary of the disc, some lattice squares are included and some are excluded. When reconstructing $f(\cdot, \cdot)$, this gives rise to an additional error, called geometric error, see Chapter 3. This error can be quantified by using the celebrated problem in analytic number theory referred to as lattice points of the circle [104, 106].

In the following we shall work with a discretized version of the Zernike moments, since we observe the function $f(\cdot, \cdot)$ in model (6.4) only on the discrete grid of points (x_i, y_j), $i, j = 1, \ldots, n$. Hence, consider the weights

$$w_{pq}(x_i, y_j) = \iint_{\Pi_{ij}} V_{pq}^*(x, y)\, dxdy, \tag{6.12}$$

where $\Pi_{ij} = [x_i - \Delta/2, x_i + \Delta/2] \times [y_j - \Delta/2, y_j + \Delta/2]$ denotes the pixel centered at (x_i, y_j). Another, even simpler version of the weights is

$$w_{pq}(x_i, y_j) = \Delta^2 V_{pq}^*(x_i, y_j). \tag{6.13}$$

The expansion (6.8) can be used to construct a truncated series estimator for $f(\cdot, \cdot)$ in model (6.4). To this end estimate the Zernike moment $A_{pq}(f)$ by

$$\widehat{A}_{pq}(f) = \sum_{(x_i, y_j) \in D} w_{pq}(x_i, y_j) Z_{i,j}, \tag{6.14}$$

where the weights are given by (6.12) or (6.13). A reconstruction algorithm of the image function $f(\cdot,\cdot)$ is then given by

$$\widehat{f}_N(x,y) = \sum_{p=0}^{N} \sum_{q=-p}^{p} n_p^{-1} \widehat{A}_{pq}(f) V_{pq}(x,y),$$

where N is a smoothing parameter which determines the number of terms in the truncated Fourier series. The mean integrated square error properties of $\widehat{f}_N(x,y)$ are discussed in Chapter 4, for general information on truncated series estimators see, e.g., [95].

6.3. Testing Rotational Symmetries

In this section we discuss how to design statistical tests for verifying rotational symmetries of $f \in L^2(D)$. We consider both d-fold rotations as well as rotational invariance, i.e., the limit case $d = \infty$.

d-Fold Rotations

Let us consider a rotation r_d through an angle of π/d for $d \in \mathbb{N}$. Since $\widetilde{(r_d f)}(\rho,\theta) = \widetilde{f}(\rho, \pi/d + \theta)$, from (6.9) it easily follows that

$$A_{pq}(r_d f) = e^{j\pi q/d} A_{pq}(f). \tag{6.15}$$

Now consider the hypothesis

$$H^{r_d} : r_d f = f \tag{6.16}$$

that the image f is invariant under the rotation r_d. Expanding both sides of (6.16) into a series with respect to the Zernike basis, we see that it is equivalent to $A_{pq}(r_d f) = A_{pq}(f)$ for all admissible pairs (p,q). In view of (6.15), this is equivalent to

$$H^{r_d} : A_{pq}(f) = e^{j\pi q/d} A_{pq}(f) \quad \text{for all admissible pairs } (p,q).$$

Therefore, a natural way to test the hypothesis H^{r_d} is via the statistic

$$T_N^{r_d} = \frac{1}{4} \sum_{p=0}^{N} \sum_{q=-p}^{p} n_p^{-1} \left|1 - e^{j\pi q/d}\right|^2 \left|\widehat{A}_{pq}(f)\right|^2, \tag{6.17}$$

where the norming factor $1/4$ is used for convenience. We will study explicitly the two most important special cases, $d = 1$, i.e., rotation through π, and $d = 2$,

rotation through $\pi/2$. In principle, one could derive similar results for general rotations of the form π/d, $d \geq 3$. However, in this case, additional technical difficulties will arise since the rectangular design is no longer invariant under r_d for $d \geq 3$.

Since for $d = 1$ we have $e^{j\pi q} = (-1)^{|q|}$, the terms in (6.17) for even q (or equivalently even p) disappear. Therefore, the statistic in this case reads

$$T_N^{r_1} = \sum_{p=0, p \text{ odd}}^{N} \sum_{q=-p}^{p} n_p^{-1} \big| \widehat{A}_{pq}(f) \big|^2.$$

The following theorem presents the asymptotic distribution of the statistic $T_N^{r_1}$ under the hypothesis H^{r_1} as well as under a fixed alternative. The notation $\xrightarrow{\mathcal{L}}$ stands for the weak convergence (convergence in distribution).

Theorem 6.1. *Let the hypothesis $H^{r_1} : r_1 f = f$ hold. If $\Delta \to 0$, $N \to \infty$ such that $\Delta N^7 \to 0$, then we have that*

$$\frac{1}{\Delta^2 \sqrt{a(N)}} \big(T_N^{r_1} - \sigma^2 \Delta^2 a(N) \big) \xrightarrow{\mathcal{L}} N(0, 2\sigma^4), \tag{6.18}$$

where $N(0, \sigma^2)$ denotes the normal law with mean zero and variance σ^2 and

$$a(N) = \begin{cases} N(N+2)/4 & : \quad N \text{ even}, \\ (N+1)(N+3)/4 & : \quad N \text{ odd}. \end{cases} \tag{6.19}$$

Under a fixed alternative $r_1 f \neq f$, suppose that $f \in C^s(D)$ for $s \geq 2$. If $\Delta N^{2s+1} \to \infty$ and $N^{3/2} \Delta^{\gamma - 1} \to 0$, where $\gamma = 285/208$ is as in Lemma 6.2, we have

$$\frac{1}{\Delta} \big(T_N^{r_1} - \|f - r_1 f\|^2 / 4 \big) \xrightarrow{\mathcal{L}} N(0, \sigma^2 \|f - r_1 f\|^2). \tag{6.20}$$

For the proof of Theorem 6.1 see the Appendix where also some technical lemmas are given.

Remark 6.1. *Let us comment on the results of Theorem 6.1. First note that different rates appear under the hypothesis in (6.18) and under fixed alternatives in (6.20). This is due to the fact that $T_N^{r_1}$ is, under the hypothesis, a quadratic statistic, but under a fixed alternative an additional linear term arises which dominates the asymptotic expansion, see [34] for a discussion of similar phenomenon in the case of testing symmetry of univariate non-parametric regression functions.*

Let us also comment on the conditions for the smoothing parameter N, which determines the number of terms in the truncated Zernike series. The condition

$\Delta N^7 \to 0$, *used under the hypothesis, is rather restrictive, and is due to the approximate orthogonality of the discretized Zernike polynomials as it is established in Lemma 6.1. In contrast, the optimal choice of* N *for estimation of a smooth function* $f \in C^1(D)$, *as was shown in Chapter 4 and [180], is of order* $N \sim \Delta^{-50/146+\epsilon}$. *However, our condition can be relaxed if we assume a more accurate orthogonal design. In fact, if we have exact orthogonality, then* $\Delta N^2 \to 0$ *is sufficient for (6.18) to continue to hold. Also note that under the hypothesis, we make no assumption on the smoothness of the image function* $f(x, y)$, *in particular it might have edges.*

Under a fixed alternative, the condition $N^{3/2}\Delta^{\gamma-1} \to 0$ *is equivalent to* $N^{4+c}\Delta \to 0$, $c = 0.0519\ldots$, *so that this condition and* $N^{2s+1}\Delta \to \infty$ *can be fulfilled simultaneously for* $s \geq 2$.

Let us now consider the problem of testing the invariance under the rotation r_2 through an angle π/d, $d = 2$. Here, the factor $\left|1 - e^{j\pi q/d}\right|^2$ in (6.17) becomes

$$\left|1 - j^q\right|^2 = \begin{cases} 4 & : \quad q \equiv 2 \bmod 4, \\ 0 & : \quad q \equiv 0 \bmod 4, \\ 2 & : \quad q \equiv 1, 3 \bmod 4. \end{cases}$$

Therefore, the statistic $T_N^{r_2}$ can be written as

$$T_N^{r_2} = \frac{1}{2} \sum_{p=0, p \text{ odd}}^{N} \sum_{q=-p}^{p} n_p^{-1} \left|\widehat{A}_{pq}(f)\right|^2 + \sum_{p=0}^{N} \sum_{q \equiv 2 \bmod 4} n_p^{-1} \left|\widehat{A}_{pq}(f)\right|^2.$$

The following theorem gives the asymptotic distribution of the test statistic $T_N^{r_2}$ both under the hypothesis $r_2 f = f$ and the fixed alternative $r_2 f \neq f$.

Theorem 6.2. *Let the hypothesis* $H^{r_2} : r_2 f = f$ *hold. If* $\Delta \to 0$, $N \to \infty$ *such that* $\Delta N^7 \to 0$, *then we have that*

$$\frac{1}{\Delta^2 \sqrt{a(N)/4 + b(N)}} \Big(T_N^{r_2} - \sigma^2 \Delta^2 \big(a(N)/2 + b(N) \big) \Big) \xrightarrow{\mathcal{L}} N\big(0, 2\sigma^4\big),$$

where $a(N)$ *is given in (6.19), and* $b(N)$ *is equal to the number of admissible* (p, q) *with* $q \equiv 2$ *mod* 4 *and* $p \leq N$.

Under a fixed alternative $r_2 f \neq f$, *suppose that* $f \in C^s(D)$ *for* $s \geq 2$. *If* $\Delta N^{2s+1} \to \infty$ *and* $N^{3/2}\Delta^{\gamma-1} \to 0$, *where* $\gamma = 285/208$ *is as in Lemma 6.2, we have that*

$$\frac{1}{\Delta} \big(T_N^{r_2} - \|f - r_2 f\|^2/4 \big) \xrightarrow{\mathcal{L}} N(0, \sigma^2 \|f - r_2 f\|^2).$$

The proof of Theorem 6.2 is very similar to that of Theorem 6.1 and is therefore omitted.

Let us, however, point to the fact that the design (x_i, y_j), $i, j = 1, \ldots, n$, is also invariant under the rotation r_2. Note that in the statistic $T_N^{r_2}$, more Zernike coefficients occur than in $T_N^{r_1}$, which is to be expected since H^{r_2} imposes more restrictions than H^{r_1}.

Implementation Issues

Theorems 6.1 and 6.2 can be used to construct an asymptotic level α test for the hypotheses H^{r_1} and H^{r_2}, respectively. Let us discuss H^{r_1}, the discussion for H^{r_2} being completely analogous.

It is easy to see that the hypothesis H^{r_1} is rejected if

$$T_N^{r_1} > u_{1-\alpha}\Delta^2\sqrt{2a(N)}\widehat{\sigma}^2 + \Delta^2 a(N)\widehat{\sigma}^2. \tag{6.21}$$

Here $u_{1-\alpha}$ denotes the $1-\alpha$ quantile of the standard normal distribution, and $\widehat{\sigma}^2$ is an estimate of the variance σ^2. To get $\widehat{\sigma}^2$, one can use a difference based estimator

$$\widehat{\sigma}^2 = \frac{1}{C(\Delta)} \sum_{(x_i,y_j)\in D} \frac{1}{4}\Big(\big(Z_{i,j} - Z_{i+1,j}\big)^2 + \big(Z_{i,j} - Z_{i,j+1}\big)^2 \Big),$$

where the sum is taken over all $(x_i, y_j) \in D$ where $(x_{i+1}, y_j) \in D$ and $(x_i, y_{j+1}) \in D$, and $C(\Delta)$ is the number of terms in this restricted sum. One can show that if the image function f is Lipschitz continuous, then $\widehat{\sigma}^2 - \sigma^2 = O_P\big(\Delta\big)$. In this case, (6.18) continues to hold if σ^2 is replaced by the estimator $\widehat{\sigma}^2$, and (6.21) is indeed an asymptotically valid level α test decision. For detailed information on difference-based estimators in higher dimensions see [162].

The asymptotic distribution (6.20) of the test statistic $T_n^{r_1}$ under a fixed alternative can be used in various ways. One is to estimate the power of the test. In fact, we have that for $\beta \in (0, 1)$,

$$P\big(\widehat{\sigma}\|f - r_1 f\|\Delta u_{1-\beta} + \|f - r_1 f\|^2/4 \le T_N^{r_1}\big) \approx \beta.$$

Using the decision rule (6.21), we see that for the power β we get asymptotically

$$\beta = 1 - \Phi\left(\frac{\widehat{\sigma}\Delta(u_{1-\alpha}\sqrt{2a(N)} + a(N))}{\|f - r_1 f\|} - \frac{\|f - r_1 f\|}{4\widehat{\sigma}\Delta} \right),$$

where $\Phi(\cdot)$ is the distribution function of the standard normal random variable. Observing that the first term in brackets tends to zero and estimating $\|f - r_1 f\|$

by $2\sqrt{T_N^{r_1}}$, we get as an estimate for the power β

$$\widehat{\beta} = 1 - \Phi\left(-\frac{\sqrt{T_N^{r_1}}}{2\widehat{\sigma}\Delta}\right).$$

Another use of (6.20) is to validate the symmetry of the image f under r_1 by testing the hypothesis

$$H_t : \|f - r_1 f\| > t \quad \text{against} \quad K_t : \|f - r_1 f\| \leq t,$$

for $t > 0$. See [35] for further details on such testing problems.

Testing Radiality

Now let us consider the problem of building a statistical test for verifying rotational invariance of f, i.e., whether f is a function of the radius ρ only, i.e., $\widehat{f}(\rho, \theta) = g(\rho)$ for some univariate function $g(\rho)$. Expressed in terms of the Zernike polynomials, a function $f \in L^2(D)$ is rotationally invariant if and only if

$$A_{pq}(f) = 0 \text{ for every } q \neq 0. \tag{6.22}$$

This is easily deduced from the definition of the Zernike coefficients (6.9), see Chapter 4. The orthogonal projection inv f of a function $f \in L^2(D)$ onto the space of rotationally invariant functions is therefore given by

$$\text{inv } f(x, y) = \sum_{p=0,\, p \text{ even}}^{\infty} n_p^{-1} A_{p,0}(f) V_{p,0}(x, y),$$

and the L^2 distance between f and inv f is

$$\|f - \text{inv } f\|^2 = \sum_{p=0}^{\infty} \sum_{q=-p, q\neq 0}^{p} n_p^{-1} |A_{pq}(f)|^2. \tag{6.23}$$

Consider the hypothesis H^{inv} that the function f is rotationally invariant,

$$H^{\text{inv}} : \text{inv } f = f.$$

Then, in order to test H^{inv}, it is natural to consider the statistic

$$T_N^{\text{inv}} = \sum_{p=1}^{N} \sum_{q=-p, q\neq 0}^{p} n_p^{-1} \left|\widehat{A}_{pq}(f)\right|^2.$$

The asymptotic distribution of T_N^{inv} is presented in the following theorem.

Theorem 6.3. *Let the hypothesis* H^{inv} *hold. If* $\Delta \to 0$, $N \to \infty$ *such that* $\Delta N^7 \to 0$, *then we have*

$$\frac{1}{\Delta^2\sqrt{a(N)}}\big(T_N^{\text{inv}} - \sigma^2\Delta^2 a(N)\big) \xrightarrow{\mathcal{L}} N(0, 2\sigma^4), \tag{6.24}$$

where

$$a(N) = \begin{cases} (N^2+2N)/2 & : \quad N \text{ even}, \\ (N+1)^2/2 & : \quad N \text{ odd}. \end{cases}$$

Under a fixed alternative $\text{inv}\, f \neq f$, *suppose that* $f \in C^s(D)$ *for* $s \geq 2$. *If* $\Delta N^{2s+1} \to \infty$ *and* $N^{3/2}\Delta^{\gamma-1} \to 0$, *where* $\gamma = 285/208$ *is as in Lemma 6.2, we have that*

$$\frac{1}{\Delta}\big(T_N^{\text{inv}} - \|f - \text{inv}\, f\|^2\big) \xrightarrow{\mathcal{L}} N(0, 4\sigma^2\|f - \text{inv}\, f\|^2). \tag{6.25}$$

The proof of this theorem is similar to that of Theorem 6.1 and is therefore omitted. Testing procedures based on Theorem 6.3 can now be implemented in a completely analogous fashion as discussed above.

General Rotations

In this section we discussed tests for invariance under specific rotations, namely d-fold rotations and rotational invariance. However, these are essentially all possibilities for rotational invariance in two dimensions. Let us consider a general rational rotation through an angle of $d_1\pi/d_2$ for $d_1, d_2 \geq 1$ coprime. Then the condition of invariance of an image under the rotation $d_1\pi/d_2$ is equivalent to invariance under π/d_2, since both reflections generate the same finite groups of rotations. As for an irrational rotation, say through $x\pi$, x irrational, it is well known that the orbit of any point on a circle of radius ρ is dense on this circle. Hence, invariance under an irrational rotation is very close to rotational invariance (for continuous images f, it is equivalent).

6.4. Testing Reflection and Joint Symmetries

Reflections

In this section we discuss how to test $f \in L_2(D)$ in model (6.4) for symmetries with respect to certain reflections. First let us consider the reflection τ at the y-axis, defined by $\tau f(x,y) = f(-x,y)$, $(x,y) \in D$. In polar coordinates, this is equivalent to

$$\widetilde{(\tau f)}(\rho,\theta) = f\big(\rho\cos(\pi-\theta), \rho\sin(\pi-\theta)\big) = \widetilde{f}(\rho, \pi-\theta).$$

A simple computation shows that $c_q(\rho, \tau f) = (-1)^{|q|} c_{-q}(\rho, f)$ and that $R_{pq}(\rho) = R_{p,-q}(\rho)$, therefore in view of (6.9) we have

$$A_{pq}(\tau f) = (-1)^{|q|} A_{p,-q}(f). \tag{6.26}$$

Now consider the hypothesis that the image f is invariant under the reflection τ, i.e.,

$$H^\tau : \tau f = f,$$

which using (6.26) can be expressed in terms of Zernike moments as $A_{pq}(f) = (-1)^{|q|} A_{p,-q}(f)$ for admissible (p, q). Therefore a natural test statistic is

$$T_N^\tau = \sum_{p=0}^{N} \sum_{q=-p}^{p} n_p^{-1} \big| \widehat{A}_{pq}(f) + (-1)^{|q|+1} \widehat{A}_{p,-q}(f) \big|^2.$$

The following theorem gives the asymptotic distribution of T_N^τ under the hypothesis H^τ as well as under a fixed alternative.

Theorem 6.4. *Let the hypothesis $H^\tau : \tau f = f$ hold. If $\Delta \to 0$, $N \to \infty$ such that $\Delta N^7 \to 0$, then we have*

$$\frac{1}{\Delta^2 \sqrt{(N+1)(N+2)}} \big(T_N^\tau - \sigma^2 \Delta^2 (N+1)(N+2) \big) \xrightarrow{\mathcal{L}} N(0, 8\,\sigma^4). \tag{6.27}$$

Under a fixed alternative $\tau f \neq f$, suppose that $f \in C^s(D)$ for some $s \geq 2$. If $\Delta N^{2s+1} \to \infty$ and $N^{3/2} \Delta^{\gamma - 1} \to 0$, where $\gamma = 285/208$ is as in Lemma 6.2, we have that

$$\frac{1}{\Delta} \big(T_N^\tau - \|f - \tau f\|^2 \big) \xrightarrow{\mathcal{L}} N(0, 16\sigma^2 \|f - \tau f\|^2). \tag{6.28}$$

The proof of this theorem is somehow similar to that of Theorem 6.1 and is therefore omitted.

Analogous test statistics and asymptotic results as in Theorem 6.4 can in principle be deduced for reflections with respect to arbitrary axis through the origin. In particular, for reflection with respect to the x-axis or one of the diagonals, similar results hold true. This is due to the fact that the design is also invariant under these reflections.

Joint Symmetries

In this section we wish to consider an important practical problem of testing joint symmetries, i.e., testing symmetry with respect to several transformations. For

example, consider testing the invariance under the reflections with respect to the x-axis and y-axis. These transformations are denoted by τ_x and τ_y, respectively.

Since $\tau_x\tau_y = r_1$ and $r_1\tau_y = \tau_x$, the group generated by $\{\tau_x, \tau_y\}$ is the same as that generated by $\{\tau_y, r_1\}$. Therefore, one can also design tests for invariance with respect to the reflection τ_y and the rotation r_1. In general, a group generated by two reflections can always be generated by a reflection and a rotation. Since we already have test statistics for the hypotheses H^{r_1} and H^{τ_y}, we could test the joint hypothesis $H^{\tau_y,r_1} = H^{r_1} \wedge H^{\tau_y}$ based on the theory of multiple testing procedures, see [100] for an introduction to the theory of multiple testing.

For example, the Bonferroni procedure tests separately both hypotheses H^{r_1} and H^{τ_y} to the level $\alpha/2$. If at least one is rejected at this level, the compound hypothesis H^{τ_y,r_1} can be rejected at a level α.

This two-step approach, however, is not optimal and it is important to construct a test statistic which directly tests the composite hypothesis H^{τ_y,r_1}. In fact, such tests often outperform multiple testing procedures in terms of power. To this end let $L_2(\tau_y, r_1) \subset L_2(D)$ be the subspace of images in $L_2(D)$ invariant under τ_y and r_1. Then the orthogonal projection π_{τ_y,r_1} onto $L_2(\tau_y, r_1)$ is given by

$$\pi_{\tau_y,r_1}(f)(x,y) = \sum_{p \text{ even}} \sum_{q=-p}^{p} n_p^{-1} \left(\frac{A_{pq}(f) + (-1)^{|q|} A_{p,-q}(f)}{2} \right) V_{pq}(x,y).$$

The test statistic can now be easily defined by estimating the distance $\|f - \pi_{\tau_y,r_1} f\|^2$. Expressing this in terms of Zernike moments leads to the following test statistic

$$\begin{aligned} T_N^{\tau_y,r_1} &= \sum_{p=0,p \text{ odd}}^{N} \sum_{q=-p}^{p} n_p^{-1} \left| \widehat{A}_{pq}(f) \right|^2 \\ &+ \frac{1}{4} \sum_{p=0,p \text{ even}}^{N} \sum_{q=-p}^{p} n_p^{-1} \left| \widehat{A}_{pq}(f) + (-1)^{|q|+1} \widehat{A}_{p,-q}(f) \right|^2. \end{aligned}$$

For this test statistic we have the following asymptotic result.

Theorem 6.5. *Let the compound hypothesis $H^{\tau_y,r_1} : \tau_y f = f$ and $r_1 f = f$ hold. If $\Delta \to 0$, $N \to \infty$ such that $\Delta N^7 \to 0$, then we have that*

$$\frac{1}{\Delta^2 \sqrt{a(N)}} \left(T_N^{\tau_y,r_1} - \sigma^2 \Delta^2 a(N) \right) \xrightarrow{\mathcal{L}} N(0, 2\sigma^4),$$

where

$$a(N) = \begin{cases} \dfrac{(N+2)^2}{8} + \dfrac{N(N+2)}{4} & N \text{ even}, \\ \dfrac{(N+1)^2}{8} + \dfrac{(N+1)(N+3)}{4} & N \text{ odd}. \end{cases}$$

Under a fixed alternative $\tau_y f \neq f$ *or* $r_1 f \neq f$ *suppose that* $f \in C^s(D)$ *for* $s \geq 2$. *If* $\Delta N^{2s+1} \to \infty$ *and* $N^{3/2}\Delta^{\gamma-1} \to 0$, *where* $\gamma = 285/208$ *is as in Lemma 6.2, we have that*

$$\frac{1}{\Delta}\big(T_N^{\tau_y,r_1} - \|f - \pi_{\tau_y,r_1} f\|^2\big) \xrightarrow{\mathcal{L}} N(0, 4\sigma^2 \|f - \pi_{\tau_y,r_1} f\|^2).$$

The proof of this theorem is somehow similar to that of Theorem 6.1 and is therefore omitted.

Testing procedures based on the result of Theorem 6.5 can now be implemented in a completely analogous fashion as in Section 6.3.

6.5. Conclusions

We have examined the basic class of symmetries which are common in natural images, i.e., rotational and reflection symmetries. We have mostly been focusing on testing an individual type of symmetry. In Section 6.4, however, a testing procedure for verifying several reflections has been developed. Generally, a challenging problem arises of testing a compound hypothesis on several possible symmetries occurring simultaneously. This requires advanced tools to be provided by the theory of multiple-hypothesis and this topic deserves further studies.

Yet there are other possible symmetries appearing in visual objects. For instance, in Fig. 6.3 the image of the yin and yang is shown which originates from ancient Chinese philosophy and metaphysics and which describes two primal opposing but complementary forces found in all things in the universe. Yin, the darker element, is passive, dark, feminine, downward-seeking, and corresponds to the night; yang, the brighter element, is active, light, masculine, upward-seeking and corresponds to the day. This image is a nonlinear counterpart of a skew-symmetric function for which $f(y,x) = -f(x,y)$, where the minus sign corresponds to the operation of changing the image color. The latter describes the skew symmetry with respect to the line $y = x$, whereas the image of the yin and yang is skew symmetric with respect to a nonlinear boundary.

6.6. Appendix

Before proving Theorem 6.1 we need two auxiliary results. The first result gives an approximation error of the discrete version of the orthogonality property of

Figure 6.3: The image of the yin and yang

the Zernike functions.

Lemma 6.1. *Let for some admissible pairs* (p, q), (p', q'),

$$I(\Delta) = \sum_{(x_i, y_j) \in D} w^*_{pq}(x_i, y_j) w_{p'q'}(x_i, y_j),$$

where $w_{pq}(x_i, y_j)$ *is defined either in (6.12) or in (6.13). Then we have for some* $c_1, c_2 > 0$ *that*

$$I(\Delta) = \Delta^2 n_p \delta_{pp'} \delta_{qq'} + c_1 \Delta^{3+\alpha} + c_2 \big(\sqrt{p + |q|} + \sqrt{p' + |q'|}\big) \Delta^{5/2}, \tag{6.29}$$

where α *can be selected as* $\alpha = 77/208 = 0.37019\ldots$.

The proof of this lemma, being somehow technical, is based on the results obtained in Chapter 4. The factor α describes the contribution of the geometric error examined extensively in Chapter 3.

The following lemma describes the discretization error of Parseval's formula when the Zernike moment $A_{pq}(f)$ is estimated by $\widehat{A}_{pq}(f)$.

Lemma 6.2. *Let* $f \in C^s(D)$ *be the class of functions having* s *continuous partial derivatives. Let* $\widehat{A}_{pq}(f)$ *be the estimate of the Zernike coefficient defined in (6.14) and let*

$$S_N = \sum_{p=0}^{N} \sum_{|q| \le p} n_p^{-1} \big| E \widehat{A}_{pq}(f) \big|^2.$$

Then we have

$$S_N = \|f\|^2 + O\big(N\Delta^{3/2} + N^{3/2}\Delta^{\gamma} + N^{-(2s+1)}\big),$$

where $\gamma = 285/208$ *describes the contribution of the geometric error, see Chapter 3 and [106].*

Remark 6.2. *Let $f \in \mathrm{BV}(D)$ be the class of functions of bounded variation on D. Then the result in Lemma 6.2reads as follows, see Chapter 4 and [180],*

$$S_N = \|f\|^2 + O\big(N\Delta^{3/2} + N^{3/2}\Delta^{\gamma} + N^{-1}\big).$$

Proof of Theorem 6.1. For admissible (p,q) with odd q, we have for the weights in both (6.12) and (6.13) that $w_{pq}(x_i,y_j) = (-1)^{|q|} w_{pq}(x_{n-i+1}, y_{n-j+1})$. This is evident for (6.13). As for (6.12), first note that the rotated version $-\Pi_{ij}$ of the pixel Π_{ij} is again a pixel, namely the pixel $\Pi_{n-i+1,n-j+1}$. Now use the fact that $w_{pq}(x_i,y_j) = A_{pq}\big(1_{\Pi_{ij}}\big)$ and (6.15). Then it follows that for admissible (p,q) with odd p,

$$\widehat{A}_{pq}(f) = \sum_{(x_i,y_j)\in D} w_{pq}(x_i,y_j)\big(\overline{f}(x_i,y_j) + \epsilon_{i,j}\big),$$

where

$$\overline{f}(x,y) = \big(f(x,y) - f(-x,-y)\big)/2.$$

Set

$$m_{(ij),(kl)} = \sum_{(p,q)} n_p^{-1} w_{pq}(x_i,y_j)\, w^*_{pq}(x_k,y_l),$$

where here and for the rest of the proof all sums involving (p,q)'s are taken over admissible pairs with $0 \le p \le N$ and odd p. We obtain

$$\begin{aligned} T_N^{r_1} &= \sum_{(x_i,y_j),(x_k,y_l)\in D} \epsilon_{i,j}\,\epsilon_{k,l}\, m_{(ij),(kl)} + 2 \sum_{(x_i,y_j),(x_k,y_l)\in D} \overline{f}(x_i,y_j)\,\epsilon_{k,l}\, m_{(ij),(kl)} \\ &\quad + \sum_{(x_i,y_j),(x_k,y_l)\in D} \overline{f}(x_i,y_j)\,\overline{f}(x_k,y_l)\, m_{(ij),(kl)} \\ &= S_{1,N} + S_{2,N} + S_{3,N}. \end{aligned} \tag{6.30}$$

First assume that the hypothesis H^{r_1} is true. Then we have that $S_{2,N} = S_{3,N} = 0$, and we only have to study $S_{1,N}$. Note that for the vectors $v_{pq} = \big(w_{pq}(x_i,y_j)\big)_{(x_i,y_j)\in D}$, from Lemma 6.1 it follows that

$$v_{pq}^{T*} v_{p'q'} = \Delta^2 n_p \delta_{p,p'}\delta_{q,q'} + O\big(\Delta^{5/2}(p^{1/2} + |q|^{1/2} + p'^{1/2} + |q'|^{1/2})\big), \tag{6.31}$$

where the constant in $O(\cdot)$ is independent of p,q,p',q'.

The matrix of coefficients $M = \big(m_{(ij),(kl)}\big)_{(i,j),(k,l)\in D}$ can be written as

$$M = \sum_{(p,q)} n_p^{-1} v_{pq}\, v_{pq}^{T*}. \tag{6.32}$$

Then using (6.31) we have for the expectation

$$\begin{aligned} ES_{1,N} = \sigma^2 trM &= \sigma^2 \sum_{(p,q)} n_p^{-1} tr(v_{pq}^{T*} v_{pq}) \\ &= \sigma^2 \sum_{(p,q)} n_p^{-1} \big(\Delta^2 n_p + O\big(\Delta^{5/2}(p^{1/2} + |q|^{1/2})\big)\big) \\ &= \sigma^2 \Delta^2 a(N) + O\big(\Delta^{5/2} N^{7/2}\big). \end{aligned} \tag{6.33}$$

Here trA denotes the trace of a square matrix A, and $a(N)$ is the number (defined in (6.19)) of terms in the sum $\sum_{(p,q)}$ with p being odd. Next decompose

$$\begin{aligned} S_{1,N} &= \sum_{(x_i,y_j)\in D} \epsilon_{i,j}^2 \, m_{(i,j),(i,j)} + \sum_{(x_i,y_j)\neq(x_k,y_l)\in D} \epsilon_{i,j}\epsilon_{k,l} m_{(ij),(kl)} \\ &= S_{1,1,N} + S_{1,2,N}. \end{aligned}$$

First note that

$$ES_{1,2,N} = ES_{1,2,N} S_{1,1,N} = 0, \qquad ES_{1,N} = ES_{1,1,N},$$

where the formula for $ES_{1,N}$ is given in (6.33).

Using the fact that $|w_{pq}(x,y)| \leq \Delta^2$, we can estimate the variance of $S_{1,1,N}$ as follows

$$\begin{aligned} \text{Var } S_{1,1,N} \leq\ & E\big(\epsilon_{1,1}^2 - \sigma^2\big)^2 \sum_{(p_1,q_1)} \sum_{(p_2,q_2)} n_{p_1}^{-1} \, n_{p_2}^{-1} \\ & \sum_{(x_i,y_j)\in D} |w_{p_1 q_1}(x_i,y_j)|^2 |w_{p_2 q_2}(x_i,y_j)|^2 = O\big(\Delta^6 N^6\big). \end{aligned} \tag{6.34}$$

Now let us evaluate the variance of $S_{1,2,N}$. Let D_M denote the diagonal matrix consisting of the diagonal elements of M. Then

$$\begin{aligned} \text{Var } S_{1,2,N} = 2\sigma^4 tr(M - D_M)^2 &= 2\sigma^4 \big(trM^2 - trD_M^2\big) \\ &= 2\sigma^4 trM^2 + O(\Delta^6 N^6), \end{aligned} \tag{6.35}$$

where we used (6.34) in the last step.

Furthermore a tedious algebra shows that

$$\begin{aligned} trM^2 &= \sum_{(p_1,q_1)} \sum_{(p_2,q_2)} n_{p_1}^{-1}\, n_{p_2}^{-1} |v^{T*}_{p_1,q_1} v_{p_2,q_2}|^2 \\ &= \sum_{(p_1,q_1)} \sum_{(p_2,q_2)} n_{p_1}^{-1}\, n_{p_2}^{-1} \big(\Delta^2 n_p \delta_{p_1,p_2}\delta_{q_1,q_2} \\ &\quad + O\big(\Delta^{5/2}(p_1^{1/2} + |q_1|^{1/2} + p_2^{1/2} + |q_2|^{1/2})\big)\big)^2 \\ &= \sum_{(p,q)} \Delta^4 + \sum_{(p,q)} n_p^{-1} \Delta^{9/2} O\big(p^{1/2} + |q|^{1/2}\big) \\ &\quad + \sum_{(p_1,q_1)} \sum_{(p_2,q_2)} n_{p_1}^{-1}\, n_{p_2}^{-1} \Delta^5 O\big(p_1 + |q_1| + p_2 + |q_2|\big) \\ &= a(N)\Delta^4 + O(\Delta^{9/2} N^{7/2}) + O(\Delta^5 N^7). \end{aligned}$$

Using this, (6.35) and $\Delta N^5 \to 0$, we finally get

$$\text{Var } S_{1,2,N} \sim 2\sigma^4 \Delta^4 a(N). \tag{6.36}$$

This is of higher order than the variance of $S_{1,1,N}$, therefore $S_{1,2,N}$ dominates the asymptotics.

To evaluate its asymptotic distribution, we use Theorem 5.2 of de Jong [32]. To check Condition 1 of this theorem, we compute

$$\begin{aligned} &\max_{(x_i,y_j)\in D} \sum_{(x_k,y_l)\in D} |m_{(ij),(kl)}|^2 \\ &= \max_{(x_i,y_j)\in D} \sum_{(p_1,q_1)} \sum_{(p_2,q_2)} n_{p_1}^{-1}\, n_{p_2}^{-1} w_{p_1q_1}(x_i,y_j) w^*_{p_2q_2}(x_i,y_j) \\ &\qquad\qquad \sum_{(x_k,y_l)\in D} w^*_{p_1q_1}(x_k,y_l) w_{p_2q_2}(x_k,y_l) \\ &= O(\Delta^6 N^6) = o\big(\Delta^4 N^2\big). \end{aligned}$$

Then Condition 2 of Theorem 5.2 in [32] will be automatically satisfied (one could choose $K(\Delta) = N$ in Condition 2 of [32]) since the sequence $\{\epsilon_{i,j}\}$ is i.i.d.

In order to check Condition 3 in [32] we have to bound the spectral value of the matrix M defined in (6.32). First note that

$$M^2 = \sum_{(p_1,q_1)} \sum_{(p_2,q_2)} n_{p_1}^{-1} n_{p_2}^{-1} v_{p_1q_1} (v^{*T}_{p_1q_1} v_{p_2q_2}) v^{*T}_{p_2q_2}.$$

Applying (6.31) yields

$$M^2 = \Delta^2 M + \Delta^{5/2} \sum_{(p_1,q_1)} \sum_{(p_2,q_2)} n_{p_1}^{-1} n_{p_2}^{-1} v_{p_1q_1} v_{p_2q_2}^{*,T} O\big(\sqrt{p_1} + \sqrt{q_1} + \sqrt{p_2} + \sqrt{q_2}\big), \quad (6.37)$$

where the summation is taken over $(p_1, q_1) \neq (p_2, q_2)$.

Let λ be an eigenvalue of the symmetric matrix M corresponding to the unit length eigenvector u. Noting that $M^2 u = \lambda^2 u$, we have from (6.37) that

$$\lambda^2 u = \Delta^2 \lambda u + \Delta^{5/2} \sum_{(p_1,q_1)} \sum_{(p_2,q_2)} n_{p_1}^{-1} n_{p_2}^{-1} v_{p_1q_1} v_{p_2q_2}^{*,T} u O\big(\sqrt{p_1} + \sqrt{q_1} + \sqrt{p_2} + \sqrt{q_2}\big).$$

Subtracting and taking the norm gives

$$\lambda^2 - \Delta^2 \lambda = \Delta^{5/2} O\left(\sum_{(p_1,q_1)} \sum_{(p_2,q_2)} n_{p_1}^{-1} n_{p_2}^{-1} \|v_{p_1q_1} v_{p_2q_2}^{*,T}\| \big(\sqrt{p_1} + \sqrt{q_1} + \sqrt{p_2} + \sqrt{q_2}\big) \right). \quad (6.38)$$

From (6.31) we obtain

$$\|v_{pq}\|^2 = n_p \Delta^2 + O\big(\Delta^{5/2}(\sqrt{p} + \sqrt{|q|})\big).$$

Thus

$$\begin{aligned} \|v_{p_1q_1} v_{p_2q_2}^{*T}\| &\leq \|v_{p_1q_1}\| \|v_{p_2q_2}\| \\ &= O\Big(\sqrt{n_{p_1} n_{p_2}} \Delta^2 + \sqrt{n_{p_1}} \big(p_2^{1/4} + |q_2^{1/4}|\big) \Delta^{9/4} \\ &\quad + \sqrt{n_{p_2}} \big(p_1^{1/4} + |q_1^{1/4}|\big) \Delta^{9/4} + \big(p_1^{1/4} + |q_1^{1/4}|\big)\big(p_2^{1/4} + |q_2^{1/4}|\big) \Delta^{5/2} \Big). \end{aligned}$$

Using this bound in (6.38), after some tedious but straightforward algebra we obtain the formula

$$\begin{aligned} \lambda(\lambda - \Delta^2) &= \Delta^{5/2} O(\Delta^2 N^{11/2} + \Delta^{9/4} N^{25/4} + \Delta^{5/2} N^7) \\ &= O(\Delta^{9/2} N^{11/2} + \Delta^{19/4} N^{25/4} + \Delta^5 N^7). \end{aligned} \quad (6.39)$$

By solving the above quadratic equation, we get

$$|\lambda| = \Delta^2 + O\big(\Delta^{9/4} N^{11/4}\big).$$

Since we have already shown that Var $S_{1,N} = O(\Delta^4 N^2)$, Condition 3 of de Jong [32] is evaluated as follows

$$\frac{|\lambda|}{\sqrt{\text{Var } S_{1,N}}} = \frac{1}{N} + O(\Delta^{1/4} N^{7/4}).$$

Since $\Delta N^7 \to 0$, and all estimates are uniform over the eigenvalues of M, this finishes the proof of the first part of Theorem 6.1.

Now let us consider the case of an alternative hypothesis, i.e., $r_1 f \neq f$. Let us start with the non-stochastic term $S_{3,N}$ in (6.30). First note that

$$S_{3,N} = \sum_{(p,q)} n_p^{-1} \big| E\widehat{A}(\overline{f}) \big|^2.$$

Then by virtue of Lemma 6.2 with $f(x,y)$ replaced by $\overline{f}(x,y)$, we have

$$S_{3,N} = \|f - r_1 f\|^2/4 + O\big(N\Delta^{3/2} + N^{3/2}\Delta^{\gamma} + N^{-(2s+1)}\big). \tag{6.40}$$

Next let us evaluate the variance of $S_{2,N}$. Note first that

$$\text{Var } S_{2,N} = 4\sigma^2 \sum_{(x_i,y_j)\in D} \left(\sum_{(x_k,y_l)\in D} \overline{f}(x_k,y_l) m_{(ij),(kl)} \right)^2.$$

By expanding the formula in brackets and recalling the definition of the matrix M we obtain

$$\text{Var } S_{2,N} = 4\sigma^2 \overline{f}^T M^2 \overline{f}, \tag{6.41}$$

where the vector $\overline{f}$ is defined by $\overline{f} = \big(\overline{f}(x_i,y_j)\big)_{(x_i,y_j)\in D}$. The proof of the first part of the theorem, see formula (6.39), reveals that

$$M^2 = \Delta^2 M + O(\Delta^{9/2} N^{11/2}).$$

This and (6.41) gives

$$\text{Var } S_{2,N} = 4\sigma^2 \Delta^2 \overline{f}^T M \overline{f} + \overline{f}^T \overline{f}\, O\big(\Delta^{9/2} N^{11/2}\big).$$

Observing that $\overline{f}^T M \overline{f} = S_{3,N}$ and using (6.40) we obtain

$$\text{Var } S_{2,N} = \sigma^2 \Delta^2 \|f - r_1 f\|^2 + O\big(N\Delta^{7/2} + N^{3/2}\Delta^{\gamma+2} + \Delta^2 N^{-(2s+1)}\big) + \|\overline{f}\|^2 O\big(\Delta^{9/2} N^{11/2}\big).$$

Hence Var $S_{2,N}$ is of order Δ^2, and $S_{2,N}$ dominates the quadratic term $S_{3,N}$. Furthermore the second term in (6.40) is negligible, even after dividing by the standard deviation Δ.

Finally, we check Lyapunov's condition. We shall show that

$$\frac{E\epsilon_{11}^4 \sum_{(x_k,y_l)\in D} \Big| \sum_{(x_i,y_j)\in D} \overline{f}(x_i,y_j) m_{(ij),(kl)} \Big|^4}{\big(\text{ Var } S_{2,N}\big)^2} \to 0. \tag{6.42}$$

The interior sum in the denominator can be evaluated as follows (see Chapter 4 and [180])

$$\begin{aligned} \sum_{(x_i,y_j)\in D} \overline{f}(x_i,y_j) m_{(ij),(kl)} &= \sum_{(p,q)} w_{pq}^*(x_k,y_l)\, n_p^{-1} \sum_{(x_i,y_j)\in D} \overline{f}(x_i,y_j) w_{pq}(x_i,y_j) \\ &= \sum_{(p,q)} w_{pq}^*(x_k,y_l)\, n_p^{-1} \big(A_{pq}(\overline{f}) + O(\Delta^\gamma)\big). \end{aligned} \tag{6.43}$$

From the proof of Lemma 4.3 in Chapter 4, see also [180], we have

$$|A_{pq}(\overline{f})| = O\left(\frac{1}{(|q|+1)(p+1)\sqrt{p-|q|+1}}\right).$$

Using this we estimate the numerator in (6.42) by

$$\sum_{(x_k,y_l)\in D} \left| \sum_{(x_i,y_j)\in D} \overline{f}(x_i,y_j) m_{(ij),(kl)} \right|^4 = O\big((N\log(N))^4\Delta^6 + \Delta^{8\gamma+6}N^{12}\big),$$

and since Var $S_{2,N}$ is of order Δ^2, (6.42) is $O((N\log(N))^4\Delta^2 + \Delta^{8\gamma+2}N^{12})$, which tends to zero since $N^{3/2}\Delta^{\gamma-1} \to 0$. This completes the proof of Theorem 6.1. $\square$

Chapter 7

Orthogonal Radial Moments for Data Hiding

This chapter concerns the application of the orthogonal radial moments to the multimedia problem of data hiding. We have already examined the invariant properties of the orthogonal radial moments and found that the invariance holds only approximately due to the inaccuracies in their computation. In this chapter we identify a subset of radial moments which are suitable for the problem of data hiding in the presence of geometric attacks. In applications of data hiding, the embedded signal robustness to geometric transformations and lossy compression is often considered to be crucial, though difficult to achieve. We present a data hiding algorithm based on the radial Zernike and pseudo-Zernike moments. By distinguishing between hardly-invariant Zernike/pseudo-Zernike and nearly-invariant Zernike/pseudo-Zernike moments, we select only the latter for data hiding to guarantee its geometric robustness. For data embedding, dither modulation is employed to quantize the magnitudes of the selected Zernike/pseudo-Zernike moments. For data extraction, a minimum distance decoder is applied to estimate the embedded data. Simulation results show that the embedded information can be extracted at low error rates, robust against typical geometric manipulations such as rotation, flipping, scaling, and aspect ratio change, as well as a variety of other distortions such as lossy compression, additive noise and lowpass filtering.

7.1. Introduction

Digital watermarking is the process of embedding some hidden information in a digital medium, such as an image, an audio signal and a video signal, by

modifying the medium slightly, and later extracting the embedded information from the modified medium for some purposes. The embedded signal is called a *watermark*, and the original medium is called the *host signal* or *cover signal*, while the modified medium is termed a *watermarked signal*.

Digital watermarking has attracted considerable attention in both academic and industry communities in the past decade, because it is regarded as a promising means to address problems related to digital media in a variety of ways. The primary application of digital watermarking is copyright protection. The drastic development of new technologies, such as the advent of the Internet and recordable compact disk, makes it extremely easy to disseminate and duplicate multimedia contents, which brings about copyright infringement. Under such circumstances, the producer of a multimedia work can use a watermark as a proof to claim authorship or ownership. Another important application of digital watermarking is authentication of multimedia content. With the wide spread use of various multimedia editing software, it is a simple matter to modify the contents of a multimedia work. A watermark can be used as an effective means to verify the integrity of a work. A digital watermark can also be useful in tracing the source of illegal copies of multimedia works. Furthermore, a digital watermark can be applied for other purposes such as device control and signal multiplexing. As an example, a typical watermarking process is shown in Fig. 7.1, where the cover signal is an image and the message to be embedded is a logo. A secret key is often used in a watermarking system for security reasons. A watermarked signal often undergoes attacks, either intentional or unintentional, before it is used for watermark extraction. Because of the attacks, the extracted message is likely to contain some errors.

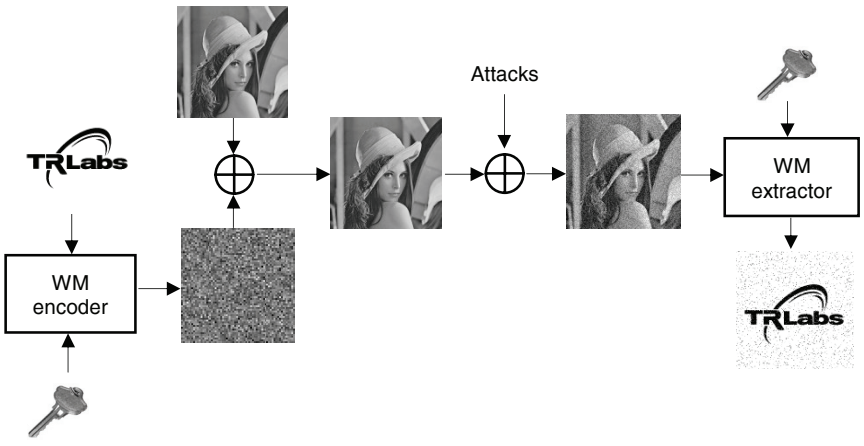


Figure 7.1: Illustration of the watermarking process

The fundamental task in designing a watermarking system is to address some conflicting requirements including watermark transparency, watermark robustness and information capacity etc [28]. Among various problems occurring in image watermarking, robustness against geometric transformations such as image rotation and scaling is considered to be the most challenging one. It is a well known fact [28] that many existing image watermarking algorithms are vulnerable to this kind of attacks. There have been some preliminary approaches in addressing this issue [4–6, 43, 123, 148, 197, 229, 239]. Among the various techniques available in the literature, we are most interested in the invariant aspects of watermarking, which are based on the invariance property of some image features. O'Ruanaidh *et al.* [165] first reported an elegant approach to the design of rotation-scaling-translation (RST) invariant watermarks in the Fourier-Mellin domain. First, discrete Fourier transform (DFT) is applied to get the translation-invariant magnitudes of Fourier coefficients, which is followed by a log-polar transform plus another DFT, resulting in scaling and rotation invariant magnitudes of Fourier-Mellin coefficients. Based on a similar idea, Lin *et al.* [148] designed a 1-D watermark whose detection statistic is invariant to image translation and scaling, but non-invariant to image rotation. A known issue associated with log-polar-based algorithms is that the image quality usually suffers greatly from the inherent instability of the inverse log-polar mapping. Yet another direction based on moment invariants has been explored for image watermarking in Alghoniemy *et al.* [5, 6]. They have utilized a set of geometric moment-based invariants, due originally to Hu [101], see also Chapter 1, to design a zero-bit watermark, which is robust against geometric manipulations and filtering, etc. The discussion in Chapter 1 points to a number of limitations of the geometric moment invariants. First, there are too few of them, which makes the embedding of a multibit watermark virtually impossible. Moreover, the nonlinear nature of the invariants makes them potentially unstable and they are ill-posed in terms of the image reconstruction.

The Zernike moments were reported as watermark features by Farzam *et al.* [43] to achieve watermark robustness to rotation, additive noise and JPEG compression. More recently, in [123], a zero-bit watermark has been embedded by modifying a Zernike feature vector consisting of some Zernike moments with orders less than five. Nevertheless, the proposed Zernike moment-based watermarking algorithms have employed poor approximates of the Zernike coefficients yielding an imperfect invariance property.

In this chapter, following our analysis of the uneven invariance property of Zernike/pseudo-Zernike moments (ZMs/PZMs), we focus on the investigation of a data hiding scheme with good robustness to geometric distortions and other

common attacks. Unlike other watermarking algorithms dealing with zero-bit watermarks, our algorithm works with multibit watermarks. This is realized by quantizing the magnitudes of the selected ZMs/PZMs through quantization index modulation [23] to embed an array of bits. We modify the magnitudes of high order ZMs/PZMs, in an attempt to embed a large data payload and simultaneously, to achieve watermark robustness to geometric distortions. In particular, rotation, scaling and flipping, as well as some other forms of attacks such as additive noise and lossy compression are taken into consideration. We obtain a watermarked image by combining the reconstructed components from the modified moments and those left intact. In watermark extraction, the same group of selected ZMs/PZMs are computed from the possibly distorted (attacked) image, and the array of bits is retrieved from the invariant magnitudes of the ZMs/PZMs using a minimum distance decoder.

This chapter is organized as follows. In Section 7.2, some preliminaries are given, including the introduction of Zernike/pseudo-Zernike polynomials and corresponding moments (ZMs/PZMs) and we discuss their invariance properties. The imperfect rotational invariance of ZMs/PZMs in the case of digital images is analyzed in Section 7.3. Subsequently we describe our watermark embedding and extraction algorithms in Section 7.4 and Section 7.5 respectively. In Section 7.6 we present a variety of simulation results which illustrate the effectiveness of the proposed algorithm. We discuss some additional aspects related to our algorithm in Section 7.7. Finally, conclusions are given in Section 7.8.

7.2. Zernike and Pseudo-Zernike Moments

Zernike/Pseudo-Zernike Polynomials

Let us recall that the Zernike basis is a set of complete and orthogonal functions on the unit disk D defined by the following rotationally invariant form [18, 250]

$$V_{nm}(x, y) = R_{nm}(\rho)e^{jm\theta}, \tag{7.1}$$

where $\rho = \sqrt{x^2 + y^2}$, $\theta = \tan^{-1}(y/x)$. Here n is a non-negative integer and m is an integer that takes positive, negative, or zero values such that $n - |m|$ is even and non-negative. The radial Zernike polynomial $R_{nm}(\rho)$ is defined by the following formula

$$R_{nm}(\rho) = \sum_{s=0}^{(n-|m|)/2} \frac{(-1)^s (n-s)! \rho^{n-2s}}{s! \left(\frac{n+|m|}{2} - s\right)! \left(\frac{n-|m|}{2} - s\right)!}. \tag{7.2}$$

A modified version of Zernike basis is the so-called pseudo-Zernike basis [13, 221], which is also a set of complete and orthogonal functions on D, and has the same invariant form as (7.1) with two exceptions. One is that m is only restricted to be $|m| \leq n$, and the other is that the radial polynomial $R_{nm}(\rho)$ is defined differently as the following

$$R_{nm}(\rho) = \sum_{s=0}^{n-|m|} \frac{(-1)^s (2n+1-s)!\, \rho^{n-s}}{s!\,(n+|m|+1-s)!\,(n-|m|-s)!}. \tag{7.3}$$

Zernike/Pseudo-Zernike Moments

Like any other orthogonal and complete basis, the Zernike/pseudo-Zernike basis can be used to decompose an analog image function $f(x, y)$:

$$f(x,y) = \sum_{n=0}^{\infty} \sum_{m} A_{nm} V_{nm}(x,y), \tag{7.4}$$

where the second summation is taken with respect to all permissible values of m. In (7.4) A_{nm} is the Zernike/pseudo-Zernike moment of order n with repetition m, which is defined as follows

$$A_{nm} = \frac{n+1}{\pi} \iint_D f(x,y) V_{nm}^*(x,y) dxdy, \tag{7.5}$$

where * denotes complex conjugate. Let us note that for the simplicity of future notation we have slightly modified the definition of A_{nm} by incorporating the normalizing factor $(n+1)/\pi$.

For digital images, (7.5) cannot be applied directly. Let us assume an $M \times M$ image matrix $F(i,j), i = 1, ..., M, j = 1, ..., M$. In order to compute its Zernike moments, one has to map $F(i,j)$ into the function $f(x_i, y_j)$ defined on $[-1,1]^2$, such that $f(x_i, y_j) = F(i,j), i = 1, ..., M, j = 1, ..., M$, where $x_i = (2i - M - 1)/M$ and $y_j = (2j - M - 1)/M$. The Zernike moment of $f(x_i, y_j)$ is computed as follows

$$\widehat{A}_{nm} = \frac{n+1}{\pi} \sum_{i=1}^{M} \sum_{j=1}^{M} h_{nm}(x_i, y_j) f(x_i, y_j), \tag{7.6}$$

where the values of i and j are taken such that $x_i^2 + y_j^2 \leq 1$, and

$$h_{nm}(x_i, y_j) = \int_{x_i - \delta/2}^{x_i + \delta/2} \int_{y_j - \delta/2}^{y_j + \delta/2} V_{nm}^*(x,y) dxdy, \tag{7.7}$$

where $\delta = 2/M$. The factor $h_{nm}(x_i, y_j)$ can be computed by using a number of numerical techniques discussed in Chapter 3. In this chapter, we adopt the following most commonly used estimate of A_{nm}

$$\widehat{A}_{nm} = \frac{n+1}{\pi}\delta^2 \sum_{i=1}^{M}\sum_{j=1}^{M} V_{nm}^*(x_i, y_j) f(x_i, y_j). \tag{7.8}$$

We have to point out that the ZMs/PZMs of discrete images computed by formula (7.8) are not very accurate. As analyzed in Chapters 3 and 4, there are two sources of errors, namely geometric and numerical integration errors. The former is due to the fact that the areas covered by the pixels involved in the computation of moments never sum up exactly to the area of the unit disk. The latter comes from the approximation of the integral defining A_{nm}.

Due to the close similarity of Zernike moments and pseudo-Zernike moments, we do not differentiate between their denotations. We denote by A_{nm} both ZMs and PZMs, and $V_{nm}(x, y)$ defines both the Zernike and pseudo-Zernike polynomials. Furthermore, in our considerations we are mostly concerned with digital images and the estimate $\widehat{A}_{nm}$ will be simply denoted by A_{nm}, unless otherwise stated.

The Invariance Properties of Zernike/Pseudo-Zernike Moments

The reason we use ZMs/PZMs for image watermarking is that they have some very important properties, i.e., their magnitudes are invariant under image rotation and image flipping. We now elaborate on these properties.

Rotation Invariance

If image $f(x, y)$ is rotated through an angle α then it can be shown that the ZMs/PZMs of the resulting image are

$$A_{nm}^{(\alpha)} = A_{nm} e^{-jm\alpha}. \tag{7.9}$$

This leads to $|A_{nm}^{(\alpha)}| = |A_{nm}|$. Therefore, if a watermark is inserted in the magnitudes of ZMs/PZMs, it is robust to rotation. Note that this property holds perfectly in the case of analog images. For the discrete version of ZMs/PZMs, typically computed by formula (7.8), this property has to be compromised to some extent. We have a detailed analysis of this problem in the next section.

Flipping Invariance

Consider another interesting property of ZMs/PZMs with respect to image flipping, either horizontal or vertical. Let the horizontally flipped version of the digital image $f(x_u, y_v)$ be defined as $f^{(hf)}(x_u, y_v) = f(-x_u, y_v)$ and the corresponding discrete ZMs/PZMs be denoted by $A_{nm}^{(hf)}$. Also let the vertically flipped version of the image $f(x_u, y_v)$ be $f^{(vf)}(x_u, y_v) = f(x_u, -y_v)$ with the corresponding discrete ZMs/PZMs $A_{nm}^{(vf)}$. Then we have

$$\begin{aligned} A_{nm}^{(hf)} &= k \sum_{\{(x_u,y_v)\in D\}} V_{nm}^*(x_u, y_v) f^{(hf)}(x_u, y_v) \\ &= k \sum_{\{(x_u,y_v)\in D\}} R_{nm}(\rho) e^{-jm\theta} f(-x_u, y_v) \\ &= k \sum_{\{(x_u,y_v)\in D\}} R_{nm}(\rho) e^{-jm(\pi-\theta)} f(x_u, y_v) \\ &= k \sum_{\{(x_u,y_v)\in D\}} (-1)^m R_{nm}(\rho) e^{jm\theta} f(x_u, y_v) \\ &= (-1)^m A_{nm}^* \end{aligned} \tag{7.10}$$

where $k = \delta^2(n+1)/\pi$.

Similarly,

$$\begin{aligned} A_{nm}^{(vf)} &= k \sum_{\{(x_u,y_v)\in D\}} V_{nm}^*(x_u, y_v) f^{(vf)}(x_u, y_v) \\ &= k \sum_{\{(x_u,y_v)\in D\}} R_{nm}(\rho) e^{-jm\theta} f(x_u, -y_v) \\ &= k \sum_{\{(x_u,y_v)\in D\}} R_{nm}(\rho) e^{-jm(-\theta)} f(x_u, y_v) \\ &= k \sum_{\{(x_u,y_v)\in D\}} R_{nm}(\rho) e^{jm\theta} f(x_u, y_v) \\ &= A_{nm}^*. \end{aligned} \tag{7.11}$$

Thus in either case, the magnitudes of the discrete ZMs/PZMs do not change, i.e., $|A_{nm}^{(hf)}| = |A_{nm}|$ and $|A_{nm}^{(vf)}| = |A_{nm}|$. This property is of significance in watermarking applications, since image flipping is an easy and effective form of attack to which many existing watermarking algorithms are vulnerable.

Scaling Invariance

In this chapter by scaling we mean image resizing. Although in theory ZMs/PZMs are not invariant to image scaling, we can still obtain approximate scaling invariance in practice, either by changing the unit disk region accordingly or resizing the image to a canonical size, provided that the unit disk is made to cover the same contents of the image. It is worth noting that this property holds well only when the image is scaled moderately. When the image is scaled to a much smaller size, this property has to be compromised due to the loss of information. However, in many applications, including the watermarking scenarios, it is true that images are scaled only slightly.

7.3. Approximate Invariance of ZMs/PZMs of Digital Images

As stated above, for digital images, we can only obtain an approximate version of ZMs/PZMs, and hence the invariance property of the moment magnitudes holds only approximately. What we are concerned with is how close this approximation is. We argue that the magnitude invariance of a particular ZM/PZM depends on its computation accuracy. We have observed that under (7.8), different moments have different levels of computation accuracy. For example, assume we have an 128×128 image of a constant graylevel 127. Based on (7.8), we have calculated its ZMs and PZMs up to order 10, whose magnitudes are shown in Table 7.1 and Table 7.2 respectively. In theory, all the ZMs/PZMs of this image except A_{00} should be zero, but in fact we can see from these tables that a number of moments deviate from zero, and some of them even have considerable magnitudes. In general, we have the following results.

Theorem 7.1. *The Zernike/pseudo-Zernike moments of a constant image $f(x_i, y_j) = T$, computed by formula (7.8), are*

$$A_{nm} = \begin{cases} T(1 + O(\delta^\gamma)), & \text{if } n = m = 0, \\ \neq 0, & \text{if } n \neq 0 \text{ and } m = 4k, k \in \mathbb{Z}, \\ 0, & \text{otherwise,} \end{cases}$$

where $1 \leq \gamma < 3/2$ is the exponent characterizing the geometric error, see Chapter 3.

This theorem is significant for invariant watermarking intended for robustness against geometric distortions. From this theorem, we know that not all the ZMs/PZMs of a discrete image can be accurately computed. The A_{nm} with

$m = 4i$, i being an integer, are not accurate and hence $|A_{nm}|$ are not invariant to image rotation. Therefore they are not suitable for invariant watermarking.

Table 7.1: Magnitudes of Zernike moments up to order 10 for a 128×128 constant image

	$m=0$	1	2	3	4	5	6	7	8	9	10
$n=0$	127.24										
1		0									
2	0.7083		0								
3		0		0							
4	1.1689		0		0.0049						
5		0		0		0					
6	1.6117		0		0.0004		0				
7		0		0		0		0			
8	2.0287		0		0.0136		0		0.2140		
9		0		0		0		0		0	
10	2.4110		0		0.0375		0		0.2169		0

Table 7.2: Magnitudes of pseudo-Zernike moments up to order 10 for a 128×128 constant image

	$m=0$	1	2	3	4	5	6	7	8	9	10
$n=0$	127.24										
1	0.4728	0									
2	0.7044	0	0								
3	0.9329	0	0	0							
4	1.1482	0	0	0	0.0049						
5	1.3682	0	0	0	0.0002	0					
6	1.5486	0	0	0	0.0080	0	0				
7	1.7667	0	0	0	0.0203	0	0	0			
8	1.8816	0	0	0	0.0377	0	0	0	0.2140		
9	2.1156	0	0	0	0.0612	0	0	0	0.1993	0	
10	2.1188	0	0	0	0.0918	0	0	0	0.1704	0	0

7.4. Watermark Embedding

The structure of the watermark embedder is depicted in Fig. 7.2. The main ideas are explained below.

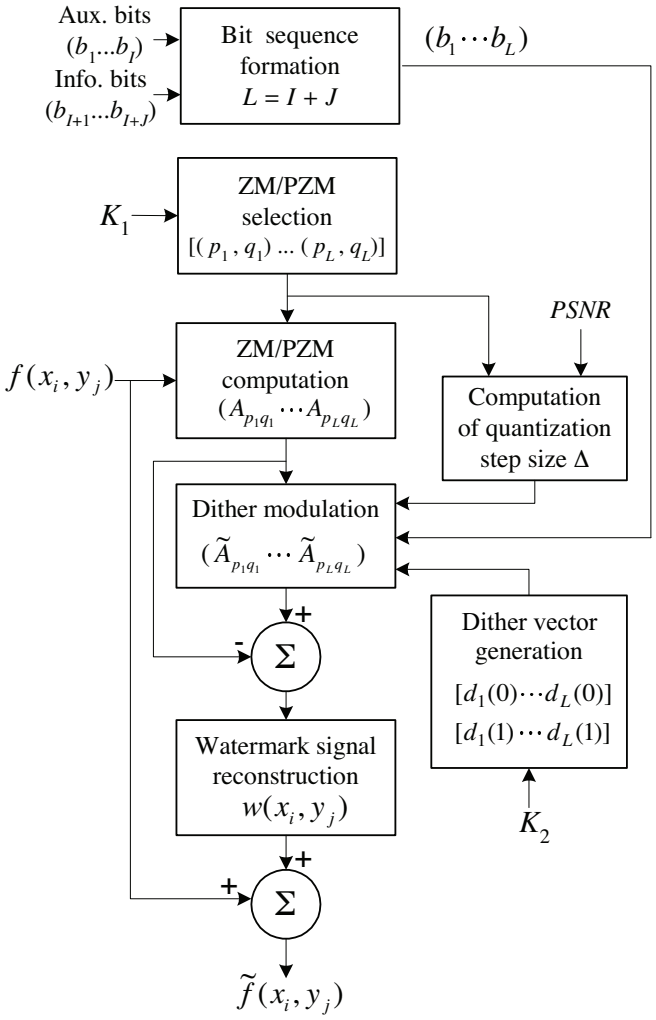


Figure 7.2: The structure of the watermark embedder

Structure of Embedded Bit Sequence

The embedded bit array consists of two parts, as indicated in Fig. 7.3. The first part is an I-bit auxiliary sequence $(b_1 ... b_I)$, which is followed by the second part, a J-bit informative sequence $(b_{I+1} ... b_{I+J})$. The auxiliary part is a fixed bit sequence, known to both the watermark embedder and watermark extractor, whose purpose is to facilitate the determination of the unit disk. Due to the possible geometric distortions, such as image resizing and scaling, it is crucial for

the watermark extractor to have the exact knowledge of the image region covered by the unit disk in order to extract the informative bits. We address this issue by embedding the fixed auxiliary bit sequence $(b_1...b_I)$. When the adopted region is not the correct unit disk region, the extracted auxiliary bit sequence $(b'_1...b'_I)$ displays randomness, and thus does not agree well with the embedded sequence $(b_1...b_I)$. On the contrary, if the unit disk region is adopted correctly, the two sequences match well. It can be shown that the probability that a uniformly distributed random bit sequence $(b'_1...b'_I)$ matches the fixed sequence $(b_1...b_I)$ at more than H bit positions is

$$P(I,H) = 2^{-I} \sum_{i=H}^{I} \binom{I}{i}. \tag{7.12}$$

For example, if the given sequence is 16 bit long, and an extracted bit sequence matches it at 14 bit positions, we are almost sure that the correct unit disk region is adopted, because $P(16,14) \approx 2.1 \times 10^{-3}$, meaning that the probability of a non-unit disk region is only 2.1×10^{-3}.

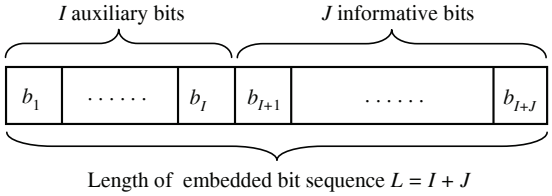


Figure 7.3: The structure of the embedded bit sequence

Selection of ZMs/PZMs

We consider three factors in selection of moments for data hiding. Firstly, as shown in the previous section, the combination of geometric error and numerical error makes the computation of some ZMs/PZMs inaccurate, thus compromising their invariance property. As a result, some ZMs/PZMs can be computed more accurately, hence more suitable for invariant data hiding than others. According to Theorem 7.1, all A_{nm}'s with repetition $m = 4i$, i being an integer, have to be ruled out for data hiding. Secondly, due to the rounding errors which become more and more significant as the order increases, there exists a certain value N_{max}, which makes a ZM/PZM A_{nm} with $n > N_{max}$ inaccurate, even if $m \neq 4i$. Therefore only those A_{nm}'s with $n \leq N_{max}$ are reliable and selected for data hiding. In our experiments, we set $N_{max} = 44$ for Zernike moments and set

$N_{max} = 23$ for pseudo-Zernike moments. Thirdly, due to the conjugate symmetry $A^*_{nm} = A_{n,-m}$, only about half of ZMs/PZMs have independent magnitudes, and in practice we only choose those A_{nm}'s with $m \geq 0$. Considering all these factors, the set of applicable ZMs/PZMs can be denoted by $\mathbb{S} = \{A_{nm}, n \leq N_{max}, m \geq 0, m \neq 4i\}$. The cardinalities of $\mathbb{S}$ can be readily obtained with straightforward algebra

$$|\mathbb{S}|_{\text{ZM}} = \begin{cases} (3N_{max}^2 + 8N_{max})/16, & \text{if } N_{max} = 4i, \\ (3N_{max}^2 + 10N_{max} + 3)/16, & \text{if } N_{max} = 4i+1, \\ (3N_{max}^2 + 8N_{max} + 4)/16, & \text{if } N_{max} = 4i+2, \\ (3N_{max}^2 + 10N_{max} + 7)/16, & \text{if } N_{max} = 4i+3, \end{cases} \tag{7.13}$$

where i is any nonnegative integer, and

$$|\mathbb{S}|_{\text{PZM}} = \begin{cases} (3N_{max}^2 + 6N_{max})/8, & \text{if } N_{max} = 4i, \\ (3N_{max}^2 + 6N_{max} - 1)/8, & \text{if } N_{max} = 4i+1, \\ (3N_{max}^2 + 6N_{max})/8, & \text{if } N_{max} = 4i+2, \\ (3N_{max}^2 + 6N_{max} + 3)/8, & \text{if } N_{max} = 4i+3, \end{cases} \tag{7.14}$$

where i is also any nonnegative integer.

Modification of ZMs/PZMs

We adopt dither modulation for the modification of ZMs/PZMs. Dither modulation is a special form of quantization index modulation, which was first proposed for data hiding in [23]. With a base quantizer $q(\cdot)$, the dither modulation function is defined as

$$f_{DM}(x; m) = q(x - d(m)) + d(m) \tag{7.15}$$

where x is a scalar variable to be quantized, m is a message symbol to be embedded in x, and $d(m)$ is the dither scalar associated with m. Dither modulation has such a property that the quantization cells and reconstruction points of any given quantizer are shifted versions of the quantization cells and reconstruction points of any other quantizer. Due to this special structure of quantizers, dither modulation has the advantage of easy implementation. In this chapter, we use binary dither modulation, i.e., $m \in \{0, 1\}$, and one independent magnitude of ZM/PZM is to carry a bit of information. In practice an independent magnitude of ZM/PZM can be used to carry more than one bits, or more than one independent magnitudes of ZMs/PZMs to carry one bit, which we will discuss later. In

this chapter a uniform scalar quantizer is adopted as our base quantizer $q(\cdot)$, i.e.,

$$q(x) = \left[\frac{x}{\Delta}\right]\Delta, \tag{7.16}$$

where $[\cdot]$ is the rounding operation, Δ is the step size of quantization.

Assume there is a bit sequence $\mathbf{b} = (b_1, ..., b_L), L \leq |\mathbb{S}|$ and $b_i \in \{0, 1\}$, to be embedded in an image $f(x_i, y_j), i, j = 1, 2, ..., M$. For the sake of security, we use a secret key K_1 to pseudorandomly choose L ZMs/PZMs from $\mathbb{S}$ to form a moment vector $\mathbf{Z} = (A_{p_1q_1}, ..., A_{p_Lq_L})$, where $A_{p_iq_i}$ is some A_{nm} in $\mathbb{S}$.

Now each bit from $\mathbf{b}$ is to be embedded into an element of $\mathbf{Z}$ via dither modulation. The magnitudes of $A_{p_iq_i}, i = 1, ..., L$ are quantized, producing a new vector $\widetilde{\mathbf{Z}} = (\widetilde{A}_{p_1q_1}, ..., \widetilde{A}_{p_Lq_L})$, where $\widetilde{A}_{p_iq_i}$ is the dither quantized version of $A_{p_iq_i}$ according to (7.15) and (7.16), satisfying

$$\begin{aligned} |\widetilde{A}_{p_iq_i}| &= f_{DM}(|A_{p_iq_i}|, b_i) \\ &= \left[\frac{|A_{p_iq_i}| - d_i(b_i)}{\Delta}\right]\Delta + d_i(b_i), \quad i = 1, ..., L \end{aligned} \tag{7.17}$$

where $d_i(\cdot)$ is the dither function for the i-th quantizer satisfying $d_i(1) = d_i(0) + \Delta/2$. The dither vector $(d_1(0), ..., d_L(0))$, whose elements are uniformly distributed over $[0, \Delta]$, is pseudorandomly generated with another key K_2, which is used to further increase the secrecy and security of the embedded signal.

The modified ZMs/PZMs are now readily calculated as

$$\widetilde{A}_{p_iq_i} = \frac{|\widetilde{A}_{p_iq_i}|}{|A_{p_iq_i}|}A_{p_iq_i}, i = 1, ..., L. \tag{7.18}$$

It is worth noting that in quantizing each $A_{p_iq_i}$, if $q_i \neq 0$, its conjugate $A_{p_i,-q_i}$ must be quantized simultaneously to ensure they always have the same magnitudes, so that the reconstructed image is real.

Determination of the Quantization Step Size

In order to use the modified ZMs/PZMs defined in (7.17) we must decide on the size of the quantization step Δ. Let us note that Δ determines the tradeoff between the visibility and robustness of a watermark. Indeed, a larger Δ gives better watermark robustness, but makes the watermark more visible. In practice, the value of Δ can be decided based on the required quality of a watermarked image. Here we use the peek signal-to-noise ratio (PSNR) defined as

$$\text{PSNR}(f, \widetilde{f}) = 10\log_{10}\frac{255^2}{\sigma_e^2}, \tag{7.19}$$

where f is the original image and $\widetilde{f}$ is the watermarked version of f, and

$$\sigma_e^2 = \frac{1}{M^2}\sum_{i=1}^{M}\sum_{j=1}^{M}[\widetilde{f}(x_i, y_j) - f(x_i, y_j)]^2 \tag{7.20}$$

is the average square error, whose relationship to Δ is expressed by the following theorem.

Theorem 7.2. *Let $\{A_{p_k q_k}\}_{k=1}^{L}$ be a selected set of Zernike/pseudo-Zernike moments of the digital image $\{f(x_i, y_j)\}$. Suppose that the set is dither-modulated by (7.17) generating a watermarked image $\{\widetilde{f}(x_i, y_j)\}$. Then assuming the high resolution quantization case the expected value of the average square error σ_e^2 defined in (7.20) is given by*

$$E\sigma_e^2 = \frac{\pi\Delta^2}{24}\sum_{k=1}^{L}(p_k + 1)^{-1} + O(M^{-2}). \tag{7.21}$$

Owing to this theorem, for a required PSNR (in dB), we can find an approximate value of the quantization step Δ as follows

$$\Delta = 255\left[10^{\frac{\mathrm{PSNR}}{10}}\frac{\pi}{24}\sum_{i=1}^{L}(p_i + 1)^{-1}\right]^{-0.5}. \tag{7.22}$$

Formation of the Watermarked Image

The reconstructed watermarked image is composed of two parts. One part is the image components contributed by the moments not selected, which is

$$f_{rem}(x_i, y_j) = f(x_i, y_j) - f_{\mathbf{Z}}(x_i, y_j), \tag{7.23}$$

where the second term is the image components contributed by the selected moments before they are changed,

$$f_{\mathbf{Z}}(x, y) = \sum_{i=1}^{L}[A_{p_i q_i}V_{p_i q_i}(x, y) + A_{p_i, -q_i}V_{p_i, -q_i}(x, y)]. \tag{7.24}$$

The other part is the image components contributed by those modified moments

$$f_{\widetilde{\mathbf{Z}}}(x, y) = \sum_{i=1}^{L}[\widetilde{A}_{p_i q_i}V_{p_i q_i}(x, y) + \widetilde{A}_{p_i, -q_i}V_{p_i, -q_i}(x, y)]. \tag{7.25}$$

Therefore we obtain a watermarked image by combining the two parts

$$\widetilde{f}(x_i, y_j) = f_{rem}(x_i, y_j) + f_{\widetilde{\mathbf{Z}}}(x_i, y_j). \tag{7.26}$$

Due to the linearity of the image reconstruction process, (7.25) can be rewritten as

$$\begin{aligned} f_{\widetilde{\mathbf{Z}}}(x,y) &= \sum_{i=1}^{L}(A_{p_iq_j} + \varepsilon_{p_iq_j})V_{p_iq_i}(x,y) + \sum_{i=1}^{L}(A_{p_i,-q_j} + \varepsilon_{p_i,-q_j})V_{p_i,-q_i}(x,y) \\ &= w(x,y) + \sum_{i=1}^{L}[A_{p_iq_i}V_{p_iq_i}(x,y) + A_{p_i,-q_i}V_{p_i,-q_i}(x,y)] \end{aligned} \tag{7.27}$$

where $\varepsilon_{p_iq_i} = \widetilde{A}_{p_iq_i} - A_{p_iq_i}$ and $\varepsilon_{p_i,-q_i} = \widetilde{A}_{p_i,-q_i} - A_{p_i,-q_i}$ are the quantization noise signals of the moments $A_{p_iq_i}$ and $A_{p_i,-q_i}$ respectively, and

$$w(x,y) = \sum_{i=1}^{L}[\varepsilon_{p_iq_i}V_{p_iq_i}(x,y) + \varepsilon_{p_i,-q_i}V_{p_i,-q_i}(x,y)] \tag{7.28}$$

is the reconstructed watermark signal, which results from the quantization noise of the selected ZMs/PZMs.

Therefore (7.26) turns into

$$\begin{aligned} \widetilde{f}(x_i, y_j) &= f_{rem}(x_i, y_j) + f_{\mathbf{Z}}(x_i, y_j) + w(x_i, y_j) \\ &= f(x_i, y_j) + w(x_i, y_j). \end{aligned} \tag{7.29}$$

7.5. Data Extraction

The process of watermark extraction is shown in Fig. 7.4. Suppose there is a test image $f'(x_i, y_j)$, which is a distorted version of $\widetilde{f}(x_i, y_j)$ after some possible manipulations, such as rotation and scaling. Our goal is to get an estimate, $\widehat{\mathbf{b}} = (\widehat{b}_1, ..., \widehat{b}_L)$, of the hidden bit sequence from $f'(x_i, y_j)$ at a low error rate. First, with the same key K_1 as in the process of watermark insertion, the identities of ZMs/PZMs involved in the data embedding process can be determined, which is denoted by $[(p_1, q_1), ..., (p_L, q_L)]$. The subsequent data extraction process can be described by the following two steps.

Locating the Unit Disk Region

The exact location of a unit disk region on the test image is crucial for the extraction of the embedded data. To facilitate the search for the unit disk region,

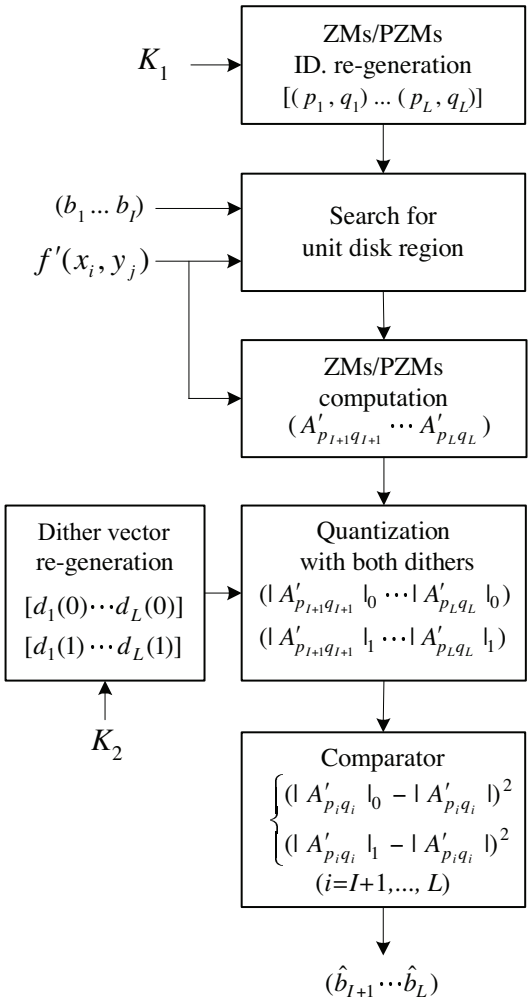


Figure 7.4: The structure of the watermark extractor

we have to use the embedded auxiliary bit sequence. For an assumed unit disk region, we extract the first I bits $(b'_1 \ldots b'_I)$, and then compare it to the known sequence $(b_1 \ldots b_I)$. As stated before, if the two sequences match well, the assumption of the unit disk region is correct. Otherwise the search process has to continue until the two sequences match well.

Informative Data Extraction

Once the unit disk region is found, we can proceed to the extraction of the informative data. First the relevant moment vector $\mathbf{Z}' = (A'_{p_{I+1}q_{I+1}}, ..., A'_{p_Lq_L})$ is computed. Then with the same key K_2 as in the embedder, the same two dither vectors $(d_1(0), ..., d_L(0))$ and $(d_1(1), ..., d_L(1))$ are re-generated. Using the same quantizer as in (7.17), we quantize the magnitude of each $A'_{p_iq_i}$ with the two corresponding dithers respectively,

$$|A'_{p_iq_i}|_j = \left[\frac{|A'_{p_iq_i}| - d_i(j)}{\Delta}\right]\Delta + d_i(j), \tag{7.30}$$

where $i = I+1, ..., L,\ j = 0, 1$, $|A'_{p_iq_i}|_j$ denotes the quantized value of $|A'_{p_iq_i}|$ with dither $d_i(j)$, and $[\cdot]$ is the rounding operation.

By comparing the distances between $|A'_{p_iq_i}|$ and its two quantized versions, we obtain the estimate of the bit embedded into $|A_{p_iq_i}|$

$$\widehat{b}_i = \arg\min_{j\in\{0,1\}}(|A'_{p_iq_i}|_j - |A'_{p_iq_i}|)^2, \quad i = I+1, ..., L. \tag{7.31}$$

which is so-called minimum distance decoder.

7.6. Simulation Results

In this section we examine the robustness of the proposed watermarking algorithm to various forms of attacks. Unless otherwise stated, the test images are 256×256 with 256 graylevels, and the unit disk is chosen such that it is fully inside an image and touches the four borders. As an example, Fig. 7.6 is the watermarked version of Fig. 7.6, in which an array of 128 bits is embedded, while Fig. 7.6 is the absolute difference between Fig. 7.6 and Fig. 7.6, multiplied by 25 for better display.

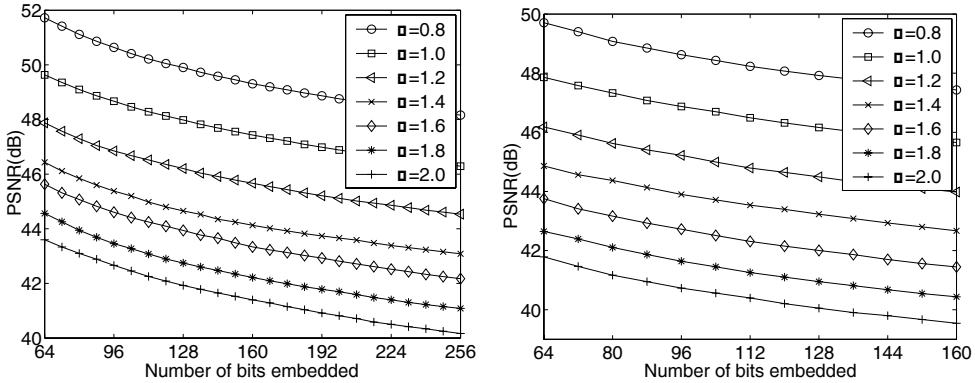
The Quality of Watermarked Images

The PSNR of a watermarked image is determined by two main factors. On the one hand, given a fixed number of bits to be embedded, the PSNR is determined by the quantization step size Δ of the dither modulation imposed on the magnitudes of ZM/PZM. A larger Δ leads to a stronger watermark, but results in a lower PSNR, and vice versa. On the other hand, given a fixed Δ or watermark strength, the number of bits to be embedded decides the PSNR of the watermarked image. The more bits embedded, the lower is the value of PSNR. The relationship between PSNR and these two factors is clearly reflected in Fig. 7.6a and Fig. 7.6b, which are the experimental results from ZM-based and PZM-based

(a) **(b)** **(c)**

Figure 7.5: An example of using the proposed algorithm.
(a) Original Lena image of size 256×256. (b) Lena watermarked with 128 bits.
(c) Exaggerated difference of (b) and (a)

algorithms respectively. Every data point in these two figures is the average of 100 individual tests. It can be verified that Fig. 7.6a and Fig. 7.6b exactly agree with Theorem 7.2.



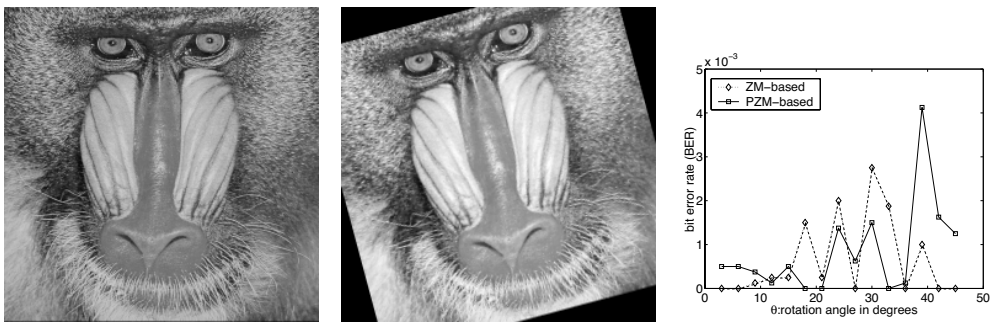
(a) **(b)**

Figure 7.6: The quality of watermarked images is affected by the number of bits embedded and the quantization step size Δ. (a) ZM-based watermarking results.
(b) PZM-based watermarking results

In all our experiments, we chose quantization step size Δ such that the resulting PSNR > 40dB, which guarantees a good watermark transparency.

Robustness to Image Rotation

We are particularly interested in the watermark performance, i.e., the bit error rate (BER), under the attack of image rotation. We used Fig. 7.7a as the test image, and 160-bit long random sequences as the information to embed. The quantization step sizes Δ were set such that an average PSNR $\approx$ 42.5 dB and 46.4 dB for ZM-based and PZM-based algorithms respectively. Fifteen different rotation angles $\theta = \{3°, 6°, ..., 45°\}$ were tried. The rotated version of each watermarked image was computed via bilinear interpolation and the resulting additional black borders were partially cropped so that the image sizes remained the same. Fig. 7.7b is an example of the rotated images. To obtain the BER at a certain rotation angle θ, 100 different randomly generated bit sequences were tried and the BER was taken as the average of the 100 cases. Fig. 7.7c shows the simulation results, both for ZM-based algorithm and the PZM-based algorithm, illustrated by the dotted line and the solid line respectively, from which we see an excellent watermark robustness to image rotation with the maximum BER at $O(10^{-3})$. It can also be seen from the figure that BER is related irregularly with θ.

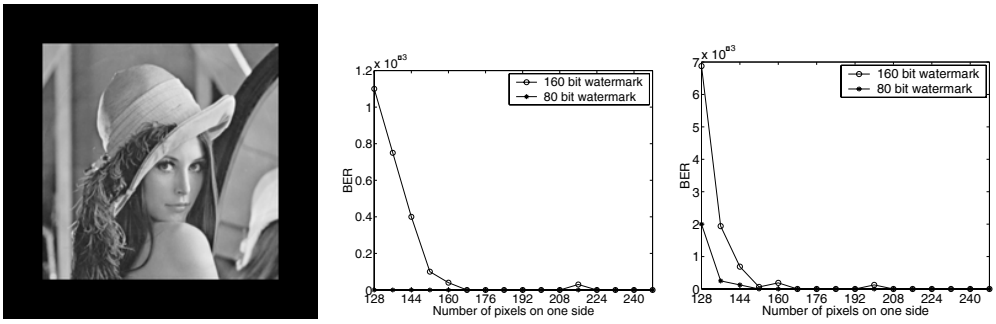


(a) **(b)** **(c)**

Figure 7.7: Watermark robustness to rotation. (a) Original baboon image of size 256×256. (b) Baboon watermarked with 160 bits followed by a 15° rotation. (c) BER as a function of rotation angles

Robustness to Image Scaling

Image scaling is another common form of geometric attacks. We looked at BERs under 16 different scaling levels. A 256×256 watermarked image was scaled to smaller sizes, ranging from 128×128 to 248×248 with an interval 8 of side length. Fig. 7.8a is an example of the scaled images. Prior to watermark ex-



(a) **(b)** **(c)**

Figure 7.8: Watermark robustness to image scaling. (a) A scaling example: watermarked Lena image scaled by 75% of side length. (b) BER as a function of scaled image size in the case of ZM-based watermarking. (c) BER as a function of scaled image size in the case of PZM-based watermarking

traction, the scaled images were scaled back to the size 256×256. It is worth noting that although they were scaled back to their original sizes, the images were quite different from their unscaled versions due to the double scaling transforms they underwent. The scaling operation was performed via bilinear interpolation. Fig. 7.8b shows the results for embedding 80 bits and 160 bits by ZM-based watermarking, while Fig. 7.8c gives the results for embedding 80 bits and 160 bits by PZM-based watermarking. Each data point in the figures is the average of 100 test results on different randomly generated bit sequences. The figure shows a trend that BER decreases as the scaling lessens and information amount drops.

Robustness to Image Flipping

Image flipping, either horizontal or vertical, is a very easy attack to perform. In fact, it is so effective that many existing watermarking algorithms perform poorly for this type of attack. Shown in Fig. 7.9a is a watermarked image. Fig. 7.9b is the horizontally flipped version of Fig. 7.9a, while Fig. 7.9c is the vertically flipped version of Fig. 7.9a. As mentioned above, the proposed approach is inherently immune, and hence perfectly robust against such an attack. In all the experiments with this kind of attack, both for ZM-based scheme and PZM-based scheme, BER = 0.

Robustness to Image Compression

With the wide-spread use of image compression standard JPEG, lossy compression is a highly common form of image processing. We looked at BERs under

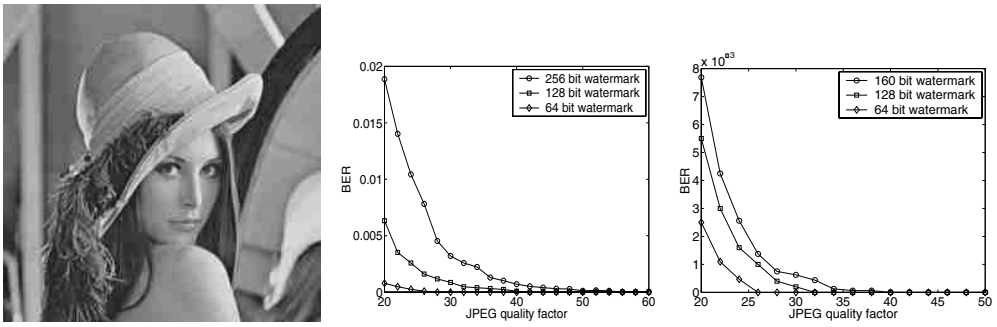
(a) **(b)** **(c)**

Figure 7.9: Watermark robustness to image flipping. (a) Watermarked image of F16. (b) Fig. (a) horizontally flipped. (c) Fig. (a) vertically flipped

different JPEG compression levels with quality factors from 20 to 90 with an interval of 2. Fig. 7.10b shows the results for embedding 64, 128 and 256 bits respectively in the image of Fig. 7.6 by means of ZM-based algorithm while Fig. 7.10c shows the results for embedding 64, 128 and 160 bits respectively in the same source image via PZM-based scheme. Each data point in the figures is the average of 100 individual results, which were obtained from 50 different randomly generated bit sequences. It can be seen that BERs decrease rapidly as the quality factor increases and the number of bits embedded drops. Considering that a JPEG quality factor less than 50 gives an obviously degraded image and hence is unlikely to be used by an attacker in practice, the robustness to JPEG lossy compression is remarkable in our algorithms.

Robustness to Lowpass Filtering

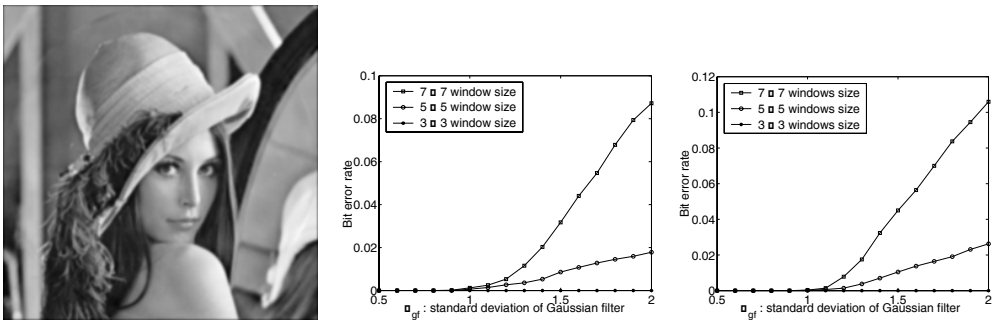
Lowpass filters are a family of filters that are commonly applied in image processing [63], including averaging filters and Gaussian filters etc., and therefore, lowpass filters are of interest to watermark designers. We introduce here the test results on the watermarked images undergoing Gaussian filtering. The test image is Lena image, the payload is 128 bits, and the average PSNR $\approx$ 41.7 dB. Shown in Fig. 7.11a is an example of Gaussian filtered watermarked image with 5×5 window and $\sigma_{gf} = 0.9$, which is apparently blurred. We recorded BERs under different levels of filter strength, σ_{gf}, ranging from 0.5 to 2 with an interval of 0.1. Fig. 7.11b shows the results for the ZM-based watermarking scheme, while Fig. 7.11c illustrates the results for the PZM-based algorithm. In both cases, 3×3, 5×5 and 7×7 window sizes of the Gaussian filters were tried. For every data point in the figures, 50 different bit sequences were generated and tested on



(a) **(b)** **(c)**

Figure 7.10: Watermark robustness to JPEG lossy compression.
(a) Watermarked Lena image compressed by JPEG with quality factor 20.
(b) BER as a function of JPEG quality factor in the case of ZM-based watermarking.
(c) BER as a function of JPEG quality factor in the case of PZM-based watermarking

the watermarked images, and then the average of 50 individual results was taken. The results display excellent watermark robustness to lowpass filtering. No error was observed when $\sigma_{gf} < 0.9$.



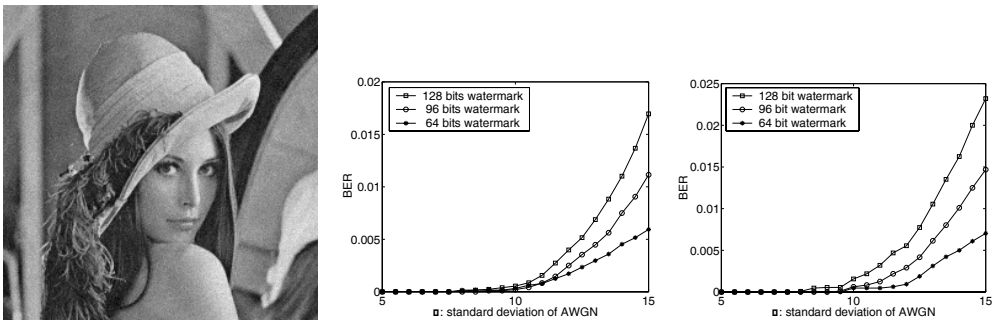
(a) **(b)** **(c)**

Figure 7.11: Watermark robustness to Gaussian filtering. (a) An example: Watermarked Lena image after a Gaussian filtering with a 5×5 window size and $\sigma_{gf} = 0.9$. (b) BER as a function of the standard deviation of the filter in the case of ZM-based watermarking. (c) BER as a function of the standard deviation of the filter in the case of PZM-based watermarking

Robustness to Additive Noise

Additive Gaussian noise is considered here to be an attack because it approximately models some interferences the watermarked images may undergo. We

recorded BERs under different levels of noise whose standard deviations range from 0 to 15 with an interval of 0.5. Fig. 7.12b and Fig. 7.12c show the results for ZM-based and PZM-based algorithms respectively, embedding 64, 96 and 128 bits in the image of Fig. 7.6, with PSNR $\approx$ 42 dB in each case. For every data point in the figures, 100 different matrices of pseudo-random Gaussian noise were tested on the watermarked image, and then the average of 100 individual results was taken. The figures show that the algorithm has an outstanding performance on the attack of additive Gaussian noise. In all our tests, we found no error when $\sigma < 5$. Even with $\sigma = 10$, BER $\leq O(10^{-4})$ if the payload is 128 bits or less. It is worth noting that when $\sigma = 5$, the attacked image displays obvious quality degradation, as shown in Fig. 7.12a, which means that an attacker has to control the noise strength such that $\sigma < 5$ in order to keep the value of the watermarked image.



(a) **(b)** **(c)**

Figure 7.12: Watermark robustness to additive Gaussian noise. (a) An example: Watermarked Lena image with Gaussian noise, $\sigma = 5$. (b) BER as a function of the standard deviation of AWGN in the case of ZM-based watermarking. (c) BER as a function of the standard deviation of AWGN in the case of PZM-based watermarking

Stirmark Test Results

Now we evaluate the watermark robustness to some of the attacks provided by the Stirmark 3.1 benchmarking tool [136]. We used the images shown in Fig. 7.13 as original images of a 512×512 size each. The unit disk was made to cover the central circular region whose diameter is 256 pixels. The auxiliary sequence used to detect the unit disk contains 16 bits, while the information to embed is 64 randomly generated bits. Each of the test images, after the 80 bits are embedded, is fed into the Stirmark tool. Then Stirmark performed various attacks on the watermarked image, and produced a series of attacked images. We used

the JPEG-compressed version of these attacked images for data extraction and recorded the number of erroneous bits for each image. The average bit error rates over all the test images are listed in Table 7.3, where -1 denotes failure to extract the embedded data. It can be seen that the proposed algorithm has excellent robustness to image rotation, scaling, cropping and JPEG compression with quality factors over 20; it has good robustness to image aspect ratio change, mild removal of lines and JPEG with very low quality factors. However, the embedded data cannot be extracted in several cases including excessive scaling, cropping, and shearing.

Table 7.3: Stirmark test results

Attack type	Average BER
Remove 17 rows and 5 columns	0.0430
Remove 5 rows and 1 columns	0.0605
Cropping 10%	0
Cropping 20%	0
Cropping 50%	0
Cropping 75%	-1
Gaussian filtering 3x3	0
JPEG 15	0.0137
JPEG 20	0.0039
JPEG 25-90	0
Change aspect ratio x:0.80 y:1.00	0.0020
Change aspect ratio x:0.90 y:1.00	0
Change aspect ratio x:1.00 y:1.10	0.0137
Rotation -0.25	0.0215
Rotation -2.00	0
Rotation 10.00	0
Rotation 90.00	0
Scale 0.25	-1
Scale 0.50	0.0020
Scale 0.75	0
Scale 0.90	0
Scale 1.50	0
Scale 2.00	0
Sharpening 3x3	0.1328
Shearing x:5.00 y:0.00	-1
Stirmark random bend	-1

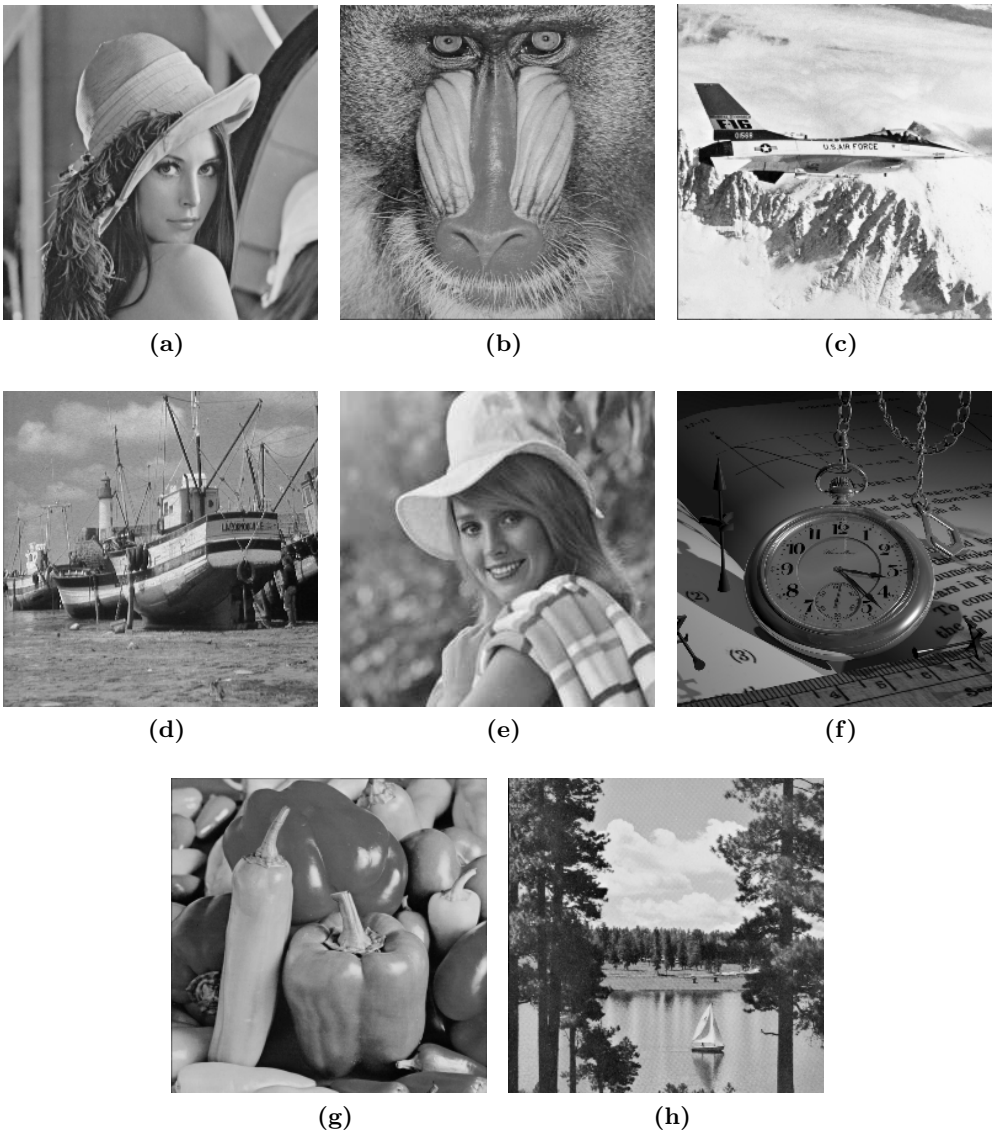


Figure 7.13: Original test images. (a) Lena. (b) Baboon. (c) F-16. (d) Fishing boat. (e) Elaine. (f) Watch. (g) Peppers. (h) Sailboat

7.7. Discussion

Comparison of ZM Watermarks with PZM Watermarks

The presented simulation results utilizing the ZM/PZM-based algorithms show that their performances are quite close in most cases. It is necessary, however, to

point out that in all the experiments, we set the quantization steps such that for the same experiment, the two algorithms would output watermarked images of the same quality. For example, in the case of 128-bit watermark, we set $\Delta = 2$ for the ZM-based algorithm while $\Delta = 1.6$ for the PZM-based algorithm, such that both algorithms produced watermarked images with PSNR ≈ 42 dB. If we use the same Δ value for the two algorithms, the PZM-based watermarks notably outperform the ZM-based watermarks in terms of robustness, but at the cost of lower quality of watermarked images. The reason for this is that for the same number of bits to be embedded, fewer lower order moments get involved in ZM-based watermarks than in PZM-based watermarks, because there are approximately twice as many PZMs as ZMs due to the constraint $n - |m| =$ even on ZMs. Roughly speaking, low order moments, which are low-frequency image components in nature, have stronger impacts on the image quality than high order moments, but have better robustness to signal distortions such as lowpass filtering and lossy compression.

The Detection of the Unit Disk Region

It is crucial that the unit disk region should be located correctly for the extraction of the embedded data. As stated before, the auxiliary bit sequence is deployed to facilitate the search for the region. In our experiments, we use finite steps of trial for the detection of the unit disk. In theory it is possible to design an algorithm to locate the unit disk in an elegant way, because the modified ZMs/PZMs of the disk region have distinguishing properties due to quantization. This interesting issue will be addressed elsewhere.

Extension to M-ary Dither Modulation

In the case of large payload, e.g., 512 bits, we can use one independent ZM/PZM magnitude to carry more than 1 bit information. The dither modulation does not have to be binary, as we have shown in the previous sections, but it can rather be M-ary in general, i.e., a set of M quantizers can be used, for which equation (7.15) still applies, where $m \in \{0, 1, ..., M-1\}$. Obviously with M-ary dither modulation, an independent magnitude of ZM/PZM carries $\log(M)$ bits of information. There is a price to pay for this gain of payload. At the same level of watermark-induced distortion, M-ary dither modulation gives higher bit error rates than binary dither modulation. This is an effective approach to the increase of data capacity of watermarks. A rigorous theoretical derivation of information capacity of ZM/PZM-based watermarks can be performed following the results obtained by some researchers, see [23].

On the other hand, in the case of small payload, more than one ZM/PZM

magnitudes can combine to carry 1 bit information in order to gain extra watermark robustness. If l independent moment magnitudes are employed to carry one bit, and binary dither modulation is performed on each of the magnitudes, then the 1-D minimum distance decoder represented by equation (7.31) would be replaced by an l-D minimum distance decoder. As a result, the watermark robustness (BER) is improved.

7.8. Conclusions

In this chapter we have illustrated the benefits of using the radial Zernike type orthogonal moments to the problem of data hiding and watermarking. Using a suitable subset of the moments, we have proposed a multibit watermarking scheme relying on the quantization-based modification of ZMs/PZMs. By quantizing the magnitudes of a group of selected ZMs/PZMs, hundreds of bits can be embedded into an image imperceptibly. An auxiliary bit sequence is used to address the location of the unit disk region in a distorted image. It has been shown that the embedded informative bit sequence can be extracted at low or even zero error rates from a distorted watermarked image. Experimental results show that the embedded data are robust against typical geometric distortions, such as image rotation, scaling, flipping, cropping and aspect ratio change, as well as other common attacks such as lossy compression, additive noise and lowpass filtering.

Chapter 8

Concluding Remarks

The method of orthogonal moments, examined in this book, provides a robust technique for representing an arbitrary image by a finite set of unique features. The technique has an appealing mathematical simplicity and is very versatile and flexible for dealing with various problems of image analysis. The fundamental property of moment descriptors is the easiness of including the concept of invariance. In fact, invariance can be achieved through the theory of algebraic invariants or can be directly embedded into moments. The latter case has been extensively examined in this book. Indeed, Chapter 3 has been concerned with a class of radial polynomials being of the invariant form. Next, in Chapter 4 we have examined the reconstruction properties of the radial invariant moments. Furthermore, in Chapter 5 we have described the efficient algorithm for accurate computation of the radial invariant moments. These findings have been extensively utilized in Chapters 6 and 7, where we have developed a statistical theory for testing image symmetry and robust algorithms for data hiding and watermarking, respectively.

The present monograph relies entirely on the original contributions of the author and his collaborators. These results have been widely published in various international journals and conference proceedings. Nevertheless, we have also presented new solutions and modified the existing theory. For example, Chapter 6 describes a novel approach to image symmetry detection combining the theory of radial invariant moments and modern statistical techniques for nonparametric lack-of-fit tests. Furthermore, Chapter 7 describes a robust and highly efficient watermarking system based on our theory. Hence, the following is a summary of the original results presented in this monograph.

- Properties of a class of orthogonal moments stemming from the theory of classical orthogonal polynomials have been thoroughly examined. This includes

issues of numerical efficiency, reconstruction power from the computed moments, robustness to noise, and automatic selection of an optimal number of moments.

- The extension of the above results to the case of generalized moments and orthogonal moments calculated in the digital domain has been given.
- Properties of radial orthogonal moments have been examined. A thorough error analysis of the moments has been conducted including numerical error, the accuracy in the presence of noise, and geometric error. In the latter case the error analysis has been performed by relating the accuracy issue to the analytic number theory of lattice point approximations.
- The reconstruction power of radial moments has been studied and optimal convergence rates for image recovery from moments have been derived.
- Fast and high precision algorithms for computing radial moments have been derived and empirically tested.
- Statistical methodology for testing image symmetry, utilizing the theory of radial moments and nonparametric statistical tests has been developed. This includes testing image rotational and reflectional symmetries as well as the compound symmetries. The limit distributions for test statistics have been established both under the null hypothesis and its fixed alternative.
- A novel watermarking system, utilizing the theory of radial moments, has been proposed. The system exhibits high robustness to geometric attacks, noise, filtering, and data compression. Watermark imbedding and extraction algorithms have been developed. Extensive empirical tests on the system accuracy have been performed.

The results of this book can form a basis for addressing other interesting problems in the theory and practice of image and shape representation. First, moment descriptors are global features and a certain form of their localization would be beneficial in recognition of occluding objects and objects with missing parts. In Chapter 2 we have briefly suggested some solution to this issue. Furthermore, in their present form, moment descriptors do not include any *a priori* knowledge about an image. We have, however, noted that radial moments can naturally incorporate some image constrains such as the radiality and symmetry. These properties have been extensively employed in Chapter 6, where the problem of testing an image symmetry has been studied.

In many applications an image is observed from various orientations and positions and collection of such images can be thought of as a manifold in the high-dimensional image space. It becomes an essential task to estimate the un-

derlying parameters (e.g., rotations) from the sequence of images. This problem has recently attracted a great deal of attention [222] and can also be approached from the point of view of the theory of moment invariants. In particular, the radial moments can serve as a natural tool for recovering the unknown rotation of an underlying set of images, see [124], [116] for some preliminary studies in this direction.

Finally, throughout the book we have assumed the white noise model. The behavior of moment descriptors in the presence of correlated noise and noise with long-range dependence would be of great interest.

Bibliography

[1] Abu-Mostafa Y.S. and Psaltis D., Recognitive aspects of moment invariants, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6:698–706, 1984.

[2] Adan A., Cerrada C., and Feliu V., Global shape invariants: a solution for 3D free-form object discrimination/identification problem, *Pattern Recognition*, 34:1331–1348, 2001.

[3] Ahmad I.A. and Li Q., Testing symmetry of an unknown density function by kernel method, *Journal of Nonparametric Statistics*, 7:279–293, 1997.

[4] Alghoniemy M. and Tewfik A.H., Geometric distortion correction through image normalization, in *IEEE Int. Conf. Multimedia and Expo*, pages 1291–1294, 2000.

[5] Alghoniemy M. and Tewfik A.H., Image watermarking by moment invariants, in *IEEE Conference on Image Processing*, pages 73–76, 2000.

[6] Alghoniemy M. and Tewfik A.H., Geometric invariants in image watermarking, *IEEE Transactions on Image Processing*, 13:145–153, 2004.

[7] Andersson R.L., Real-time gray-scale video processing using a moment-generating chip, *IEEE Journal on Image Processing of Robotics and Automation*, 1(2):79–85, 1985.

[8] Bailey R.R. and Srinath M., Orthogonal moment features for use with parametric and non-parametric classifiers, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18:389–396, 1996.

[9] Bary N.K., *A Treatise on Trigonometric Series*, Pergamon Press, Oxford, 1964.

[10] Bavinck H., On absolute convergence of Jacobi series, *Journal of Approximation Theory*, 4:387–400, 1971.

[11] Belkasim S. and Kamel M., Fast computation of 2-d image moments using biaxial transform, *Pattern Recognition*, 34:1867–1877, 2001.

[12] Bentkus V. and Götze F., Lattice point problems and distribution of values of quadratic forms, *Annals of Mathematics*, 150:977–1027, 1999.

[13] Bhatia A.B. and Wolf E., On the circle polynomials of Zernike and related orthogonal sets, *Proceedings of Cambridge Philosophical Society*, 50:40–53, 1954.

[14] Bickel P.J., On adaptive estimation, *The Annals of Statistics*, 10:647–671, 1982.

[15] Bin Y. and Jia-Xiong P., Invariance analysis of improved Zernike moments, *Journal of Optics A: Pure and Applied Optics*, 4:606–614, 2002.

[16] Blu T., Thevenaz P., and Unser M., Linear interploation revitalized, *IEEE Transactions on Image Processing*, 13:710–719, 2004.

[17] Bober M., MPEG-7 visual shape descriptors, *IEEE Transactions on Circuits and Systems for Video Technology*, 11:716–719, 2001.

[18] Born M. and Wolf E., *Principles of Optics*, Pergamon Press, Oxford, 1975.

[19] Campbell C.E., A new method for describing the aberrations of the eye using Zernike polynomials, *Optom. Vision Science*, 80(1):77–83, 2003.

[20] Casasent D., Advanced optical processors for multiple degree-of-freedom object recognition, *IEEE Transactions on Aerospace and Electronic Systems*, 24:608–618, 1988.

[21] Catmull E. and Rom R., A class of local interpolating splines, in R.E. Barnhill and R.F. Riesenfeld, editors, *Computer Aided Geometric Design*, pages 317–326, Academic Press, New York, 1974.

[22] Chang T., Spherical regression, *The Annals of Statistics*, 14:907–924, 1989.

[23] Chen B. and Wornell G.W., Quantization index modulation methods: a class of provably good methods for digital watermarking and information embedding, *IEEE Transactions on Information Theory*, 47(4):1423–1443, 2001.

[24] Chen K., Efficient parallel algorithm for the computation of two-dimensional image moments, *Pattern Recognition*, 23:109–119, 1990.

[25] Chong C.W., Raveendran P., and Mukundan R., Translation invariants of Zernike moments, *Pattern Recognition*, 36:1765–1773, 2003.

[26] Chong C.W., Raveendran P., and Mukundan R., Translation and scale invariants of Legendre moments, *Pattern Recognition*, 37:119–129, 2004.

[27] Chung K.L., Computing horizontal/vertical convex shape's moments on reconfigurable meshes, *Pattern Recognition*, 29:1713–1717, 1996.

[28] Cox I.J., Miller M.L., and Bloom J.A., *Digital Watermarking*, Morgan Kaufmann Publishers, 2001.

[29] Dai M., Baylou P., and Najim M., An efficient algorithm for computation of shape moments from run-length codes or chain codes, *Pattern Recognition*, 25(10):1119–1128, 1992.

[30] Davies P.J., Plane regions determined by complex moments, *Journal of Approximation Theory*, 19:148–153, 1977.

[31] Davis P.J. and Rabinowitz P., *Methods of Numerical Integration*, Academic Press, New York, 1975.

[32] de Jong P., A central limit theorem for generalized quadratic forms, *Probability Theory and Related Fields*, 75:261–277, 1987.

[33] Deans S., *The Radon Transform and Some of Its Applications*, Wiley, New York, 1983.

[34] Dette H., Kusi-Appiah S., and Neumeyer N., Testing symmetry in nonparametric regression models, *Journal of Nonparametric Statistics*, 14:477–494, 2002.

[35] Dette H. and Munk A., Validation of linear regression models, *The Annals of Statistics*, 14:778–800, 1998.

[36] Dodgson N., Quadratic interpolation for image resampling, *IEEE Journal on Image Processing*, 6(9):1322–1326, 1997.

[37] Donoho D.L. and Johnstone I.M., Projection-based approximation and a duality with kernel method, *The Annals of Statistics*, 17:58–106, 1989.

[38] Duda R.O., Hart P.E., and Stork D.G., *Pattern Classification*, Wiley, New York, 2002.

[39] Efromovich S., *Nonparametric Curve Estimation: Methods, Theory, and Applications*, Springer-Verlag, New York, 1999.

[40] Engels H., *Numerical Quadrature and Cubature*, Academic Press, London, 1980.

[41] Eubank R.L., *Spline Smoothing and Nonparametric Regression*, Marcel Dekker, New York, 1988.

[42] Fan Y. and Gencay R., A consistent nonparametric test of symmetry in linear regression models, *Journal of American Statistical Association*, 90:551–557, 1995.

[43] Farzam M. and Shirani S., A robust multimedia watermarking technique using Zernike transform, in *IEEE Int. Workshop Multimedia Signal Processing*, pages 529–534, 2001.

[44] Flusser J., On the independence of rotation moment invariants, *Pattern Recognition*, 33:1405–1410, 2000.

[45] Flusser J., Refined moment caculation using image block representation, *IEEE Transactions on Image Processing*, 9(11):1977–1978, 2000.

[46] Flusser J., Affine invariants of convex polygons, *IEEE Transactions on Image Processing*, 11:1117–1118, 2002.

[47] Flusser J., On the inverse problem of rotation moment invariants, *Pattern Recognition*, 35:3015–3017, 2002.

[48] Flusser J., Boldys J., and Zitova B., Invariants to convolution in arbitrary dimensions, *Journal of Mathematical Imaging and Vision*, 13:101–113, 2000.

[49] Flusser J., Boldys J., and Zitova B., Moment forms invariant to rotation and blur in arbitrary number of dimensions, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25:234–246, 2003.

[50] Flusser J. and Suk T., Pattern recognition by affine moment invariants, *Pattern Recognition*, 26:167–174, 1993.

[51] Flusser J. and Suk T., Degraded image analysis: an invariant approach, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20:590–603, 1998.

[52] Flusser J., Suk T., and Saic S., Image features invariant with respect to blur, *Pattern Recognition*, 28:1723–1732, 1995.

[53] Flusser J., Suk T., and Saic S., Recognition of blurred images by the method of moments, *IEEE Transactions on Image Processing*, 5:533–538, 1996.

[54] Forsyth D., Mundy J.L., Zisserman A., Coelho C., Heller A., and Rothwell C., Invariant descriptors for 3-d object recognition and pose, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13:971–991, 1991.

[55] Fraser W. and Gotlieb C.C., A calculation of the number of lattice points in the circle and sphere, *Mathematics of Computation*, 16:282–292, 1962.

[56] Friedberg S.A., Finding axes of skewed symmetry, *Computer Vision, Graphics, and Image Processing*, 32(2):138–155, 1986.

[57] Gautschi W., On generating orthogonal polynomials, *SIAM Journal on Scientific and Statistical Computing*, 3:289–317, September 1982.

[58] Gesu V.D. and Palenichka R.M., A fast recursive algorithm to compute local axial moments, *Signal Processing*, 81:265–273, 2001.

[59] Ginis H.S., Plainis S., and Pallikaris A., Variability of wavefront aberration measurements in small pupil sizes using a clinical Shack-Hartman aberrometer, *BMC Ophthalmology*, 4(1), 2004.

[60] Giusti E., *Minimal Surfaces and Functions of Bounded Variation*, Birkhäuser, Boston, 1984.

[61] Goldenshluger A. and Spokoiny V., On the shape-from-moments problem and recovering edges from noisy Radon data, *Probability Theory and Related Fields*, 6:176–184, 2003.

[62] Golub G.H., Milanfar P., and Varah J., A stable numerical method for inverting shape from moments, *SIAM Journal on Scientific Computing*, 21:1222–1243, 1999.

[63] Gonzalez R. and Woods R., *Digital Image Processing*, Prentice-Hall, Inc., New York, 2002.

[64] Gradshteyn I.S. and Ryzhik I.M., *Table of Integrals, Series, and Products*, Academic Press, San Diego, 5th edition, 1994.

[65] Grardner R.J. and Milanfar P., Reconstruction of convex bodies from brightness functions, *Discrete and Computational Geometry*, 29:279–303, 2003.

[66] Greblicki W., Krzyżak A., and Pawlak M., Distribution free pointwise consistency of kernel regression estimate, *The Annals of Statistics*, 12:1570–1575, 1984.

[67] Greblicki W. and Pawlak M., Classification using the Fourier series estimate of multivariate density functions, *IEEE Transactions on Systems, Man, and Cybernetics*, 11:726–730, 1981.

[68] Greblicki W. and Pawlak M., A classification procedure using the multiple Fourier series, *Information Sciences*, 26:115–126, 1982.

[69] Greblicki W. and Pawlak M., Almost sure convergence of classifying procedure using Hermite series density estimate, *Pattern Recognition Letters*, 2:13–17, 1983.

[70] Greblicki W. and Pawlak M., Hermite series estimate of a probability density and its derivatives, *Journal of Multivariate Analysis*, 15:174–182, 1984.

[71] Greblicki W. and Pawlak M., Discussion on the paper "Some aspects of the spline smoothing approach to nonparametric regression curve fitting" by B. Silverman, *Journal of Royal Statistical Society*, 47:36–37, 1985.

[72] Greblicki W. and Pawlak M., Fourier and Hermite series regression estimates, *Annals of the Institute of Statistical Mathematics*, 37:443–454, 1985.

[73] Greblicki W. and Pawlak M., Pointwise consistency of the Hermite series density estimate, *Statistics and Probability Letters*, 3:65–69, 1985.

[74] Greblicki W. and Pawlak M., Nonparametric identification of discrete Hammerstein systems, *IEEE Transactions on Automatic Control*, 31(1):74–77, 1986.

[75] Greblicki W. and Pawlak M., Hammerstein system identification by nonparametric regression estimation, *International Journal of Control*, 45:343–354, 1987.

[76] Greblicki W. and Pawlak M., Necessary and sufficient conditions for Bayes risk consistency of recursive kernel classification rule, *IEEE Transactions on Information Theory*, 33:408–412, 1987.

[77] Greblicki W. and Pawlak M., Necessary and sufficient pointwise consistency conditions for recursive kernel regression estimate, *Journal of Multivariate Analysis*, 23:67–76, 1987.

[78] Greblicki W. and Pawlak M., Nonparametric identification of Hammerstein systems, *IEEE Transactions on Information Theory*, 35:409–418, 1989.

[79] Greblicki W. and Pawlak M., Recursive nonparametric identification of Hammerstein systems, *Journal of Franklin Institute*, 326:461–481, 1989.

[80] Greblicki W. and Pawlak M., Nonparametric identification of a cascade nonlinear time series system, *Signal Processing*, 22:61–75, 1991.

[81] Greblicki W. and Pawlak M., Nonparametric identification of a particular nonlinear time series system, *IEEE Transactions on Signal Processing*, 40:985–989, 1992.

[82] Greblicki W. and Pawlak M., Dynamic system identification with order statistics, *IEEE Transactions on Information Theory*, 40(5):1475–1489, 1994.

[83] Greblicki W. and Pawlak M., Nonparametric identification of a cascade nonlinear system, *International Journal of Systems Science*, 25(1):129–153, 1994.

[84] Greblicki W. and Pawlak M., Nonparametric recovering of nonlinearities in systems with the help of Laguerre polynomials, *Control-Theory and Advanced Technology*, 10(4):771–791, 1994.

[85] Greblicki W. and Pawlak M., Discussion on the paper "Subspace identification of multivariate Hammerstein and Wiener models" by J. Gomez and E. Baeyens, *European Journal of Control*, 11:145–148, 2005.

[86] Gustafsson B., He C., Milanfar P., and Putinar M., Reconstructing planar domains from their moments, *Inverse Problems*, 16:1053–1070, 2000.

[87] Guy R.K., *Unsolved Problems in Number Theory*, chapter Gauss' lattice point problems, pages 240–241, Springer-Verlag, New York, 1994.

[88] Haddadnia J., Faez K., and Ahmadi M., A neural based human face recognition system using an efficient feature extraction method with pseudo Zernike moment, *Journal of Circuits, Systems, and Computers*, 11:283–304, 2002.

[89] Hall P., Kay J.N., and Titterington D.M., On estimation of noise variance in two-dimensional signal processing, *Advances in Applied Probability*, 23:476–495, 1991.

[90] Hall P. and Koch S., On the feasibility of cross-validation in image analysis, *SIAM Journal on Applied Mathematics*, 52(1):292–313, 1992.

[91] Hall P. and Raimondo M., On the global performance of approximations to smooth curves using gridded data, *The Annals of Statistics*, 26:2206–2217, 1998.

[92] Härdle W., Hall P., and Marron J.S., How far are automatically chosen regression smoothing parameters from their optimum, *Journal of American Statistical Association*, 83:86–101, 1988.

[93] Hardy G.J., On the double Fourier series and especially those which represent the double zeta-function with real and incommensurable parameters, *Quarterly Journal of Mathematics*, 37:53–68, 1906.

[94] Hardy G.J., Littlewood J.E., and Polya G., *Inequalities*, Cambridge University Press, Cambridge, 1952.

[95] Hart J.D., *Nonparametric Smoothing and Lack-of-Fit Tests*, Springer, 1997.

[96] Hasiewicz Z., Pawlak M., and Śliwiński P., Nonparametric identification of nonlinearities in block-oriented systems by orthogonal wavelets with compact support, *IEEE Transactions on Circuits and Systems - I*, 52:427–442, February 2005.

[97] Hastie T.J. and Tibshirani R.J., *Generalized Additive Models*, Chapman and Hall, London, 1990.

[98] Hatamian M., A real-time two-dimensional moment generating algorithm and its single chip implementation, *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 34(3):546–553, 1986.

[99] Hildebrandt T.H. and Schoenberg I.J., On linear functional operations and the moment problem for a finite interval in one or several dimensions, *Annals of Mathematics*, 34:317–328, 1933.

[100] Hsu J.C., *Multiple Comparisons – Theory and Methods*, Chapman and Hall, London, 1996.

[101] Hu M.K., Visual problem recognition by moment invariant, *IRE Transactions on Information Theory*, 8:179–187, February 1962.

[102] Hu P.H., Stone J., and Stanley T., Applications of Zernike polynomials to atmospheric propagation problems, *Journal of the Optical Society of America A*, 6:1595–1608, 1989.

[103] Hussain B. and Kabuka M.R., Real-time system for accurate three-dimensional position determination and verification, *IEEE Transactions on Robotics and Automation*, 6:31–43, 1990.

[104] Huxley M.N., Exponential sums and lattice points, *Proceedings of the London Mathematical Society*, 60:471–502, 1990.

[105] Huxley M.N., *Area, Lattice Points, and Exponential Sums*, Clarendon Press, Oxford, 1996.

[106] Huxley M.N., Integer points, exponential sums and the Riemann zeta function, in *Number Theory for the Millennium*, volume II, pages 275–290, Peters, London, 2002.

[107] Iskander D.R., Collins M.J., and Davis B., Optimal modeling of corneal surfaces with Zernike polynomials, *IEEE Transactions on Biomedical Engineering*, 48(1):87–95, 2001.

[108] Iskander D.R., Collins M.J., Morelande M.R., and Zhu M., Analyzing the dynamic wavefront aberrations in the human eye, *IEEE Transactions on Biomedical Engineering*, 51(11):1969–1980, 2004.

[109] Iskander D.R., Morelande M.R., Collins M.J., and Davis B., Modeling of corneal surfaces with radial polynomials, *IEEE Transactions on Biomedical Engineering*, 49(4):320–328, 2002.

[110] Iwaniec H. and Mozzochi C.J., On the divisor and circle problems, *Journal of Number Theory*, 29:60–93, 1988.

[111] Jacob M., Blu T., and Unser M., An exact method for computing the area moments of wavelet and spline curves, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23:633–642, 2001.

[112] Jain A.K., *Fundamentals of Digital Image Processing*, Prentice Hall, Englewood Cliffs, New York, 1989.

[113] Jaklic A. and Solina F., Moments of superellipsoids and their application to range image registration, *IEEE Transactions on Systems, Man, and Cybernetics – Part B*, 33:648–657, 2003.

[114] Jiang X.Y. and Bunke H., Simple and fast computation of moments, *Pattern Recognition*, 24(8):801–806, 1991.

[115] Jones M.C. and Silverman B.W., An orthogonal series density estimation approach to reconstructing positron emission tomography images, *Journal of Applied Statistics*, 16:177–191, 1989.

[116] Kakarala A.A. and Cadzow J., Estimation of phase for noisy linear phase signals, *IEEE Transactions on Signal Processing*, 44:2483–2497, 1996.

[117] Kamgar-Parsi B. and Kamgar-Parsi B., Evaluation of quantization error in computer vision, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(9):929–940, September 1989.

[118] Kanaya N., Iiguni Y., and Maeda H., 2-D DOA estimation method using Zernike moments, *Signal Processing*, 82:521–526, 2002.

[119] Karatsuba A.A., *Basic Analytic Number Theory*, Springer-Verlag, Berlin, 1993.

[120] Keys R.G., Cubic convolution interpolation for digital image processing, *IEEE Transactions ASSP*, 29(6):1153–1160, 1981.

[121] Khotanzad A. and Hong Y.H., Invariant image recognition by Zernike moments, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12:489–498, 1990.

[122] Khotanzad A. and Hong Y.H., Rotation invariant image recognition using features selected via a systematic method, *Pattern Recognition*, 23:1089–1101, 1990.

[123] Kim H.S. and Lee H.K., Invariant image watermark using Zernike moments, *IEEE Transactions on Circuits and Systems for Video Technology*, 13(8):766–775, 2003.

[124] Kim W.Y. and Kim Y.S., Robust rotation angle estimator, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21:768–773, 1999.

[125] Kiryath N. and Gofman Y., Detecting symmetry in grey level images: the global optimization approach, *International Journal of Computer Vision*, 29:29–45, 1998.

[126] Koornwinder T., *Two-variable analogues of the classical orthogonal polynomials*, chapter Theory and Applications of Special Functions, pages 435–495, Academic Press, New York, 1975.

[127] Korostelev A.P. and Tsybakov A.B., *Minimax Theory of Image Reconstruction*, Springer-Verlag, Berlin, 1993.

[128] Kozek A. and Pawlak M., Universal consistency of kernel nonparametric M-estimators, *Statistics and Probability Letters*, 58:343–353, 2002.

[129] Kronrod A.S., On functions of two variables, *Uspehi Math. Nauk*, 5:24–134, 1950.

[130] Krzyżak A., Leung S.Y., and Suen C.Y., Reconstruction of two-dimensional patterns from Fourier Descriptors, *Machine Vision and Applications*, 2:123–140, 1989.

[131] Krzyżak A. and Pawlak M., Universal consistency results for Wolverton-Wagner regression function estimates with applications in discrimination, *Problems of Control and Information Theory*, 12:33–42, 1983.

[132] Krzyżak A. and Pawlak M., Almost everywhere convergence of recursive regression estimate and classification, *IEEE Transactions on Information Theory*, 30:91–93, 1984.

[133] Krzyżak A. and Pawlak M., Distribution free consistency of nonparametric kernel regression estimate and classification, *IEEE Transactions on Information Theory*, 30:78–81, 1984.

[134] Krzyżak A. and Pawlak M., The pointwise rate of convergence of the kernel regression estimate, *Journal of Statistical Planning and Inference*, 16:159–166, 1987.

[135] Krzyżak A., Rafajłowicz E., and Pawlak M., Moving average algorithms for band-limited signal recovery, *IEEE Transactions on Signal Processing*, 45:2967–2976, 1997.

[136] Kutter M., Watermarking resisting to translation, rotation, and scaling, in *Proc. SPIE Multimedia Systems Applications*, pages 423–431, 1998.

[137] Landau H.J., *Moments in Mathematics*, American Mathematical Society, Providence, R.I., 1987.

[138] Leblanc F. and Lepski O.V., Test of symmetry in nonparametric regression, *Theory of Probability and its Applications*, 47:34–52, 2003.

[139] Lestrel P.E., *Fourier Descriptors and Their Applications in Biology*, Cambridge University Press, 2001.

[140] Li B. and Shen J., Range-image-based calculation of three-dimensional convex object moments, *IEEE Transactions on Robotics and Automation*, 9:484–490, 1993.

[141] Li B.C. and Shen J., Fast computation of moment invariants, *Pattern Recognition*, 24(8):807–813, 1991.

[142] Li B.C. and Shen J., Pascal triangle transform approach to the calculation of 3d moments, *Graphical Models and Image Processing*, 54:301–307, 1992.

[143] Li Y., Reforming the theory of invariant moments for pattern recognition, *Pattern Recognition*, 25:723–730, 1992.

[144] Liao S.X. and Pawlak M., On image analysis by moments, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18:254–266, 1996.

[145] Liao S.X. and Pawlak M., On the accuracy of Zernike moments for image analysis, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(12):1358–1364, 1998.

[146] Liao X. and Pawlak M., Chinese character recognition via orthogonal moments, in J.Y. Chouinard, P. Fortier, and T.A. Gulliver, editors, *Lecture Notes in Computer Science*, volume 1133, pages 296–308, Springer-Verlag, New York, 1996.

[147] Liao X. and Pawlak M., A study of Zernike moment computing, in R. Chin and T.C. Pong, editors, *Lecture Notes in Computer Science*, volume 1351, pages 394–401, Springer-Verlag, New York, 1998.

[148] Lin C.Y., Wu M., Bloom J.A., Cox I.J., Miller M.L., and Liu Y.M., Rotation, scale, and translation resilient watermarking for images, *IEEE Transactions on Image Processing*, 10(5):767–782, 2001.

[149] Lisani J.L., Moisan L., Monasse P., and Morel J.M., On the theory of planar shape, *SIAM Multiscale Modeling and Simulation*, 1:1–24, 2003.

[150] Logan B.F. and Shepp L.A., Optimal reconstruction of a function from its projections, *Duke Mathematical Journal*, 42:645–659, 1975.

[151] Lorentz G.G., *Approximation of Functions*, Holt, Rinehart and Winston, New York, 1966.

[152] Lugosi G. and Pawlak M., On the posterior probability estimate of the error rate of nonparametric classification rules, *IEEE Transactions on Information Theory*, 40(2):475–481, 1994.

[153] Maeland E., On the comparison of interpolation methods, *IEEE Transactions on Medical Imaging*, 7(3):213–217, 1988.

[154] Mallat S., *A Wavelet Tour of Signal Processing*, Academic Press, New York, 2nd edition, 1999.

[155] Marola G., On the detection of the axes of symmetry of symmetric and almost symmetric planar images, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(1):104–108, 1989.

[156] Martinez J. and Thomas F., Efficient computation of local geometric moments, *IEEE Transactions on Image Processing*, 11:1102–1111, 2002.

[157] Milanfar P., Karl W.C., and Willsky A.S., A moment-based variational approach to tomographic reconstruction, *IEEE Transactions on Image Processing*, 5:459–470, 1996.

[158] Mukundan R., Ong S.H., and Lee P.A., Image analysis by Tchebicheff moments, *IEEE Transactions on Image Processing*, 10:1357–1364, 2001.

[159] Mukundan R. and Ramakrishnan K.R., Fast computation of Legendre and Zernike moments, *Pattern Recognition*, 28(9):1433–1442, 1995.

[160] Mukundan R. and Ramakrishnan K.R., *Moment Functions in Image Analysis: Theory and Applications*, World Scientific, Singapore, 1998.

[161] Mundy J. and Zisserman A., *Geometric Invariance in Computer Vision*, The MIT Press, Boston, 1997.

[162] Munk A., Bissantz N., Wagner T., and Freitag G., On difference-based variance estimation in nonparametric regression when the convariate is high dimensional, *Journal of Royal Statistical Society B*, 67:19–41, 2005.

[163] Newey W.K. and Powell J.L., Asymmetric least squares estimation and testing, *Econometrica*, 55:819–847, 1987.

[164] Olver P.J., *Classical Invariant Theory*, Cambridge University Press, Cambridge, 1999.

[165] O'Ruanaidh J.J.K. and Pun T., Rotation, scale and translation invariant spread spectrum digital image watermarking, *Signal Processing*, 66(8):303–317, 1998.

[166] Palaniappan R., Raveendran P., and Omatu S., New invariant moments for non-uniformly scaled images, *Pattern Analysis & Applications*, 3:78–87, 2000.

[167] Pawlak M., On nonparametric estimation of a functional of a probability density, *IEEE Transactions on Information Theory*, 32(1):73–84, 1986.

[168] Pawlak M., Discussion on the paper "What is projection pursuit" by M.C. Jones, *Journal of Royal Statistical Society A*, 150:31–32, 1987.

[169] Pawlak M., On the asymptotic properties of smoothed estimators of the classification error-rate, *Pattern Recognition*, 21:515–524, 1988.

[170] Pawlak M., Note on the asymptotic properties of smoothed estimators of the classification error-rate, *Pattern Recognition*, 22:221, 1989.

[171] Pawlak M., On the almost everywhere properties of the kernel regression estimate, *Annals of the Institute of Statistical Mathematics*, 43:311–326, 1991.

[172] Pawlak M., On the series expansion approach to the identification of Hammerstein systems, *IEEE Transactions on Automatic Control*, 36:763–767, 1991.

[173] Pawlak M., On the reconstruction aspects of moment descriptors, *IEEE Transactions on Information Theory*, 38(6):1698–1708, November 1992.

[174] Pawlak M., Kernel classification rules from missing data, *IEEE Transactions on Information Theory*, 39:979–988, 1993.

[175] Pawlak M., Discontinuity estimation in nonparametric regression via orthogonal series, in J. Sall and A. Lehman, editors, *Computing Science and Statistics*, volume 26, pages 252–256, Interface Foundation of North America, 1995.

[176] Pawlak M., Nonparametric estimation of the G/G/1 queueing system, in A. Alfa and S. Chakravarthy, editors, *Advances in Matrix Analytic Methods for Stochastic Models*, pages 279–294, Notable Publications, 1998.

[177] Pawlak M., Signal sampling and recovery under dependent noise, *Journal on Sampling Theory and Image Processing*, 1:77–86, 2002.

[178] Pawlak M. and Greblicki W., Nonparametric estimation of a class of nonlinear time series models, in G.G. Roussas, editor, *NATO Advanced Study Institute, Nonparametric Functional Estimation and Related Topics*, Kluwer Academic Publishers, 1991.

[179] Pawlak M. and Hasiewicz Z., Nonlinear system identification by the Haar multiresolution analysis, *IEEE Transactions on Circuits and Systems*, 45:945–961, 1998.

[180] Pawlak M. and Liao S.X., On the recovery of a function on a circular domain, *IEEE Transactions on Information Theory*, 48(10):2736–2753, 2002.

[181] Pawlak M. and Liao X., On the digital approximation of moment descriptors, *Machine Graphics & Vision*, 3:61–68, 1994.

[182] Pawlak M. and Rafajłowicz E., On restoring band-limited signals, *IEEE Transactions on Information Theory*, 40(5):1490–1503, 1994.

[183] Pawlak M. and Rafajłowicz E., Vertically weighted regression – a tool for constructing control charts, *Statistical Archives*, 84:367–388, 2000.

[184] Pawlak M. and Rafajłowicz E., Jump preserving signal reconstruction using vertical weighting of samples, *Journal of Nonlinear Analysis*, 47:327–338, 2001.

[185] Pawlak M. and Rafajłowicz E., Non-linear harmonic filters for edge-preserving image denoising, in *The 16th International Conference on Pattern Recognition*, Quebec City, August 2002.

[186] Pawlak M., Rafajłowicz E., and Krzyżak A., Double exponential weighting algorithm for band-limited signals restoration, *IEEE Transactions on Signal Processing*, 44:538–545, 1996.

[187] Pawlak M., Rafajłowicz E., and Krzyżak A., Post-filtering versus pre-filtering for signal recovery from noisy samples, *IEEE Transactions on Information Theory*, 49:3195–3212, December 2003.

[188] Pawlak M., Rafajłowicz E., and Steland A., On detecting jumps in time series – nonparametric setting, *Journal of Nonparametric Statistics*, 16:329–347, 2004.

[189] Pawlak M. and Schmid W., On the distributional properties of ARCH models, *Journal of Time Series*, 22:339–352, 2001.

[190] Pawlak M. and Siu D., Classification with noisy features, *Advances in Pattern Recognition*, 1451:845–852, 1999.

[191] Pawlak M. and Stadtmüller U., Recovering band-limited signals under noise, *IEEE Transactions on Information Theory*, 42:1425–1438, 1996.

[192] Pawlak M. and Stadtmüller U., Kernel density estimators from quantized data, *Nonlinear Analysis Theory, Methods & Applications*, 30:3553–3559, 1997.

[193] Pawlak M. and Stadtmüller U., Kernel regression estimators for signal recovery, *Statistics and Probability Letters*, 31:185–198, 1997.

[194] Pawlak M. and Stadtmüller U., Nonparametric estimation of a class of smooth functions, *Journal of Nonparametric Statistics*, 8:149–183, 1997.

[195] Pawlak M. and Stadtmüller U., Kernel density estimation with generalized binning, *Scandinavian Journal of Statistics*, 26:539–561, 1999.

[196] Pawlak M. and Stadtmüller U., Statistical aspects of sampling for noisy and grouped data, in J. Benedetto and J. Ferreira, editors, *Modern Sampling Theory: Mathematics and Applications*, pages 317–342, Birkhauser, New York, 2001.

[197] Pereira S. and Pun T., Robust template matching for affine resistant image watermarks, *IEEE Transactions on Image Processing*, 9(6):1123–1129, 2000.

[198] Philips W., A new fast algorithm for moment computation, *Pattern Recognition*, 26:1619–1621, 1993.

[199] Prokop R.J. and Reeves A.P., A survey of moment-based techniques for unoccluded object representation and recognition, *Graphical Models and Image Processing*, 54(5):438–460, September 1992.

[200] Rafajłowicz E. and Pawlak M., On function recovery by neural networks based on orthogonal expansions, *Nonlinear Analysis Theory, Methods & Applications*, 30:1343–1354, 1997.

[201] Rafajłowicz E. and Pawlak M., Optimization of centers' positions for RBF nets with generalized kernels, in L. Rutkowski, editor, *Advances in Neural*

Computing, Lecture Notes in AI 3070, pages 253–259, Springer-Verlag, New York, 2004.

[202] Reiss T.H., The revised fundamental theorem of moment invariants, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13:830–832, 1991.

[203] Reiss T.H., *Recognizing Planar Objects Using Invariant Image Features*, Springer-Verlag, Berlin, 1993.

[204] Rothe I., Susse H., and Voss K., The method of normalization to determine invariants, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(4):366–375, 1996.

[205] Rudin L.I., Osher S., and Fatemi E., Nonlinear total variation based noise removal algorithms, *Physica D*, 60:259–268, 1992.

[206] Sansone G., *Orthogonal Functions*, Dover Publications, Inc., New York, 1991.

[207] Schwiegerling J., Greivenkamp J.E., and Miller J.M., Representation of videokeratoscopic height data with Zernike polynomials, *Journal of the Optical Society of America A*, 12:2105–2113, 1995.

[208] Shu H.Z., Luo L.M., Yu W.X., and Fu Y., A new fast method for computing Legendre moments, *Pattern Recognition*, 33:341–348, 2000.

[209] Shu H.Z., Luo L.M., Yu W.X., and Zhou J.D., Fast computation of Legendre moments of polyhedra, *Pattern Recognition*, 34:1119–1126, 2001.

[210] Simonoff J.S., *Smoothing Methods in Statistics*, Springer, New York, 1996.

[211] Slepian D., Prolate spheroidal wave functions, Fourier analysis and uncertainty-IV: Extensions to many dimensions; generalized prolate spheroidal functions, *The Bell Systems Technical Journal*, 43:3009–3056, 1964.

[212] Spiliotis I.M. and Mertzios B.G., Real-time computation of two-dimensional moments on binary images using image block representation, *IEEE Transactions on Image Processing*, 7:1609–1615, 1998.

[213] Suhling M., Arigovindan M., Hunziker P., and Unser M., Multiresolution moment filters: Theory and applications, *IEEE Transactions on Image Processing*, 13:484–495, 2004.

[214] Suk T. and Flusser J., Combined blur and affine moment invariants and their use in pattern recognition, *Pattern Recognition*, 36:2895–2907, 2003.

[215] Suk T. and Flusser J., Projective moment invariants, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26:1364–1367, 2004.

[216] Swanson M.D., Kobayashi M., and Tewfik A.H., Multimedia data-embedding and watermarking techniques, *Proceedings of the IEEE*, 86(6):1064–1087, 1998.

[217] Szegö G., *Orthogonal Polynomials*, volume 23, American Mathematical Society Colloquie Publications, Providence, R.I., 4th edition, 1975.

[218] Talenti G., Recovering a function from a finite number of moments, *Inverse Problems*, 3:501–517, 1987.

[219] Teague M.R., Image analysis via the general theory of moments, *Journal of the Optical Society of America A*, 70:920–930, August 1980.

[220] Teh C.H. and Chin R.T., On digital approximation of moment invariants, *Computer Vision, Graphics, and Image Processing*, 33:318–326, 1986.

[221] Teh C.H. and Chin R.T., On image analysis by the methods of moments, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 10:496–512, July 1988.

[222] Tenenbaum J.B., de Silva V., and Langford J.C., A global geometric framework for nonlinear dimensionality reduction, *Science*, 290:2319–2323, 2000.

[223] Thompson A.M., Brown J.C., Kay J.W., and Titterington D.M., A study of methods of choosing the smoothing parameter in image restoration by regularization, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(4):326–339, April 1991.

[224] Timan A.F., *Theory of Approximation of Functions of a Real Variable*, Dover Publications, Inc., New York, 1994.

[225] Trappe W., Wu M., Wang Z.J., and Liu K.J.R., Anti-collusion fingerprinting for multimedia, *IEEE Transactions on Signal Processing*, 51(4):1069–1087, 2003.

[226] Trier O.D., Jain A.K., and Taxt T., Feature extraction methods for character recognition – a survey, *Pattern Recognition*, 29:641–701, 1996.

[227] Tuzikov A.V., Sheynin S.A., and Vasiliev P.V., Computation of volume and surface body moments, *Pattern Recognition*, 36:2521–2529, 2003.

[228] Verral S.C. and Kakarala R., Disk-harmonic coefficients for invariant pattern recognition, *Journal of the Optical Society of America A*, 15:389–401, 1998.

[229] Voloshynovskiy S., Deguillaume F., and Pun T., Multibit digital watermarking robust against local nonlinear geometrical distortions, in *IEEE Conference on Image Processing*, pages 999–1102, 2001.

[230] Wahba G., *Spline Models for Observational Data*, SIAM Press, Philadelphia, 1990.

[231] Wallin A. and Kubler O., Complete sets of complex Zernike moment invariants and the role of the pseudoinvariants, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 17:1106–1110, 1995.

[232] Wang L. and Healey G., Using Zernike moments for the illumination and geometry invariant classification of multispectral texture, *IEEE Transactions on Image Processing*, 7:196–203, 1998.

[233] Watson A.B., Efficiency of an image code based on human vision, *Journal of the Optical Society of America A*, 4(12):2401–2417, 1987.

[234] Weiss I., Geometric invariants and object recognition, *International Journal of Computer Vision*, 10:201–231, 1993.

[235] Wilton J.R., The lattice points of a circle: an historical account of problem, *Messenger of Mathematics*, 58:67–80, 1929.

[236] Wu C.H., Horng S.J., and Lee P.Z., A new computation of shape moments via quadtree decomposition, *Pattern Recognition*, 34:1319–1330, 2001.

[237] Wu M. and Liu B., Data hiding in image and video: Part I – Fundamental issues and solutions, *IEEE Transactions on Image Processing*, 12(6):685–695, 2003.

[238] Xiao Z., Hou Z., Miao C., and Wang J., Using phase information for symmetry detection, *Pattern Recognition Letters*, 26:1985–1994, 2005.

[239] Xin Y., Pawlak M., and Liao S., A multibit geometrically robust image watermark based on Zernike moments, in *International Conference on Pattern Recognition (ICPR) 2004*, volume IV, pages 861–864, 2004.

[240] Xin Y., Pawlak M., and Liao S., A multibit geometrically robust image watermarking based Zernike moments, in *International Conference on Pattern Recognition*, pages 1051–1055, Cambridge, August 2004.

[241] Xin Y., Pawlak M., and Liao S., Image reconstruction with polar Zernike moments, in *International Conference on Advances in Pattern Recognition (ICAPR) 2005*, Bath, United Kingdom, August 2005.

[242] Xin Y., Pawlak M., and Liao S., Image reconstruction with polar Zernike moments, in S. Singh, editor, *Advances in Pattern Recognition*, volume 3687 of *Lecture Notes in Computer Science*, pages 394–403, Springer-Verlag, New York, 2005.

[243] Xin Y., Pawlak M., and Liao S., On the improvement of rotational invariance of Zernike moments, in *International Conference on Image Processing (ICIP) 2005*, Genova, Italy, September 2005.

[244] Xin Y., Pawlak M., and Liao S., Circularly orthogonal moments for robust data hiding, *IEEE Transactions on Image Processing (submitted)*, 2006.

[245] Yahampath P. and Pawlak M., On finite-state vector quantization for noisy channels, *IEEE Transactions on Communication*, 52:2125–2133, December 2004.

[246] Yang L. and Albregtsen F., Fast and exact computation of cartesian geometric moments using discrete Green's theorem, *Pattern Recognition*, 29:1061–1073, 1996.

[247] Yang Y., Hsu Y., and Arsenault H., Optimal circular symmetrical filters and their uses in optical pattern recognition, *Optica Acta*, 29:627–644, 1982.

[248] Yap P.T., Paramesran R., and Ong S.H., Image analysis by Krawtchouk moments, *IEEE Transactions on Image Processing*, 12:1367–1377, 2003.

[249] Zakaria M.F., Vroomen L.J., Zsombor-Murray P.J.A., and van Kessel J.M.H.M., Fast algorithm for the computation of moment invariants, *Pattern Recognition*, 20:639–643, 1987.

[250] Zernike F., Beugungstheorie des Schneidenverfahrens und seiner verbesserten Form, der Phasenkontrastmethode, *Physica*, 1:689–701, 1934.

[251] Zhang D. and Lu G., Review of shape representation and description techniques, *Pattern Recognition*, 37:1–19, 2004.

[252] Zhang J., Zhang X., Krim H., and Walter G.G., Object representation and recognition in shape spaces, *Pattern Recognition*, 36:1143–1154, 2003.

[253] Zhang Y., Wen C., Zhang Y., and Soh Y.C., On the choice of consistent canonical form during moment normalization, *Pattern Recognition Letters*, 24:3205–3215, 2003.

[254] Ziemer W.P., *Weakly Differentiable Functions*, Springer, New York, 1989.