A Survey of Orthogonal Moments for Image Representation: Theory, Implementation, and Evaluation

SHUREN QI, YUSHU ZHANG, and CHAO WANG, Nanjing University of Aeronautics and Astronautics, China
JIANTAO ZHOU, University of Macau, China
XIAOCHUN CAO, Sun Yat-sen University, China

Image representation is an important topic in computer vision and pattern recognition. It plays a fundamental role in a range of applications toward understanding visual contents. Moment-based image representation has been reported to be effective in satisfying the core conditions of semantic description due to its beneficial mathematical properties, especially geometric invariance and independence. This article presents a comprehensive survey of the orthogonal moments for image representation, covering recent advances in fast/accurate calculation, robustness/invariance optimization, definition extension, and application. We also create a software package for a variety of widely used orthogonal moments and evaluate such methods in a same base. The presented theory analysis, software implementation, and evaluation results can support the community, particularly in developing novel techniques and promoting real-world applications.

CCS Concepts: • **Computing methodologies** → **Computer vision representations**; • **Mathematics of computing** → **Functional analysis**; **Numerical analysis**; • **General and reference** → **Surveys and overviews;**

Additional Key Words and Phrases: Pattern recognition, image representation, orthogonal moments, geometric invariance, fast computation

ACM Reference format:
Shuren Qi, Yushu Zhang, Chao Wang, Jiantao Zhou, and Xiaochun Cao. 2021. A Survey of Orthogonal Moments for Image Representation: Theory, Implementation, and Evaluation. *ACM Comput. Surv.* 55, 1, Article 1 (November 2021), 35 pages.
https://doi.org/10.1145/3479428

This work was supported in part by the National Key R&D Program of China under Grant No. 2018AAA0100600, in part by the National Natural Science Foundation of China under Grants No. 62072237 and No. 61971476, in part by the Research Fund of Guangxi Key Lab of Multi-Source Information Mining & Security under Grant No. MIMS20-02, in part by the Guangxi Key Laboratory of Trusted Software under Grant No. KX202027, in part by the Basic Research Program of Jiangsu Province under Grant No. BK20201290, and in part by the Macau Science and Technology Development Fund under Grant No. 077/2018/A2.

Authors' addresses: S. Qi, Y. Zhang (corresponding author), and C. Wang, College of Computer Science and Technology, Nanjing University of Aeronautics and Astronautics, Jiangjun Road, Nanjing, Jiangsu, China, and Collaborative Innovation Center of Novel Software Technology and Industrialization, Nanjing University of Aeronautics and Astronautics, Jiangjun Road, Nanjing, Jiangsu, China; emails: {shurenqi, yushu, c.wang}@nuaa.edu.cn; J. Zhou, State Key Laboratory of Internet of Things for Smart City, University of Macau, Avenida da Universidade, Macau, China, and Department of Computer and Information Science, University of Macau, Avenida da Universidade, Macau, China; email: jtzhou@umac.mo; X. Cao, School of Cyber Science and Technology, Sun Yat-sen University, Gongchang Road, Shenzhen, Guangdong, China; email: caoxch5@sysu.edu.cn.

0360-0300/2021/11-ART1 $15.00
https://doi.org/10.1145/3479428

1 INTRODUCTION

In the mid-20th century, American mathematician Claude Elwood Shannon published his paper *A mathematical theory of communication* [1], which marked the creation of information theory. As a landmark contribution, information theory is the theoretical foundation of information storage, processing, and transmission in modern computer systems. On the other hand, it also dictates that the raw signal (i.e., digital data) of multimedia (e.g., image) is not semantic in nature. Therefore, one of the main requirements in many computer vision and pattern recognition applications is to have a "meaningful representation" in which semantic characteristics of digital image are readily apparent, as shown in Figure 1. This process is often termed as *image representation*. For example, for recognition, the representation should highlight the most salient semantics; for denoising, it should efficiently distinguish between signal (semantically relevant) and noise (semantically irrelevant); and for compression, it should capture the most semantic information using the least coefficients.

Mathematically, a basic idea of image representation is to project the original image function onto a space formed by a set of specially designed basis functions and obtain the corresponding feature vector. For many years, dictionary (i.e., the set of basis functions) design has been pursued by many researchers in roughly two different paths: *handcrafted* and *deep learning* [2].

Recently, deep learning techniques, represented by **Convolutional Neural Networks** (**CNNs**), have led to very good performance on a variety problems of computer vision and pattern recognition. Deep learning–based image representations are formed by the composition of multiple nonlinear transformations, mapping raw image data directly into abstract semantic representations without manual intervention (i.e., end-to-end paradigm). For such representation learning methods, the dictionary can be considered as a composite function and is trained/learned by backpropagating error. Deep learning–based image representations offer great flexibility and the ability to adapt to specific signal data. Due to the data-driven nature, this line of approaches is strongly influenced by the latest advances in optimization algorithms, computing equipment, and training data. As a result, the deep learning approaches exhibit limitations in the following three aspects [3]: (1) The quality of the representation depends heavily on the completeness of the training data, i.e., a large and diverse training set is required. (2) The time/space cost of these approaches is often very high, which prevents them from being used in time-critical applications. (3) The robustness to geometric transformation is limited, requiring the data augmentation to enhance the geometric invariance, but at the cost of time/space complexity. In contrast, handcrafted image representations are still competitive in the above three aspects. It is also worth mentioning that the successful experiences behind handcrafted features are instructive for the design of deep learning methods such as **Principal Component Analysis Network** (**PCANet**) [4] and **Spatial Pyramid Pooling** (**SPP**) [5].

In the pre-CNN era, handcrafted representations and feature engineering had made important contributions to the development of computer vision and pattern recognition. At current stage, handcrafted features still cannot be completely replaced, considering that some limitations of deep learning are just the characteristics of handcrafted representations [6]. The existing handcrafted image representation methods can be roughly divided into four categories [7]:

- *Frequency transform* – such as Fourier Transform, Walsh Hadamard Transform, and Wavelet Transform.
- *Texture* – such as **Scale Invariant Feature Transform** (**SIFT**), **Gradient Location and Orientation Histogram** (**GLOH**), and **Local Binary Patterns** (**LBPs**).
- *Dimensionality reduction* – such as **Principal Component Analysis** (**PCA**), **Singular Value Decomposition** (**SVD**), and **Locally Linear Embedding** (**LLE**).

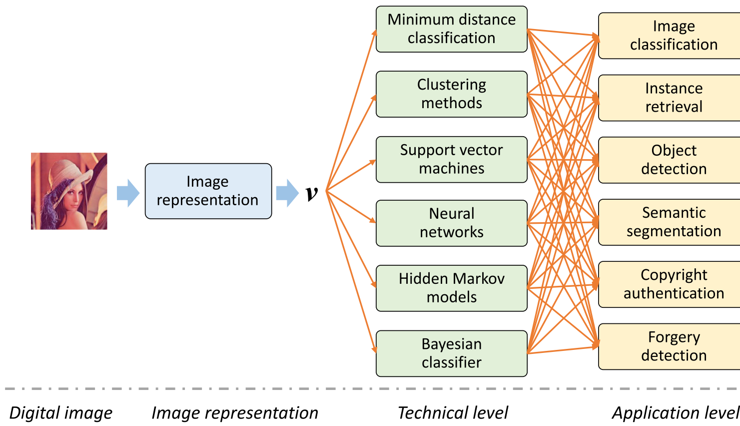


Fig. 1. The fundamental role of image representation in computer vision and pattern recognition applications. Typical visual system is generally formed on the basis of a "meaningful representation" of digital image (image representation), then followed by knowledge extraction techniques (technical level), and finally achieved the high-level visual understanding (application level).

— *Moments and moment invariants* – such as **Zernike Moments** (ZM), Legendre Moments, and **Polar Harmonic Transforms** (PHTs).

Starting from the semantic nature of the representation task, it is clear that image representation should satisfy the following basic conditions [8]:

— *Discriminability* – the representation reflects inter-class variations, i.e., two objects from two different classes have different features.
— *Robustness* – the representation is not influenced by intra-class variations, i.e., two objects from one class have the same features.

Handcrafted representations based on frequency transform, texture, and dimensionality reduction have been widely used in real-world applications. However, due to the inherent *semantic gap* between low-level descriptors and high-level visual concepts, these methods have flaws in one or both of robustness and discriminability. One example is that SIFT features exhibit *synonymy* and *polysemy* [9], caused by poor discriminability and robustness. In contrast, moments and moment invariants perform better in overcoming the semantic gap due to their beneficial mathematical properties:

— *Independence* – the orthogonality of basis functions ensures no information redundancy in moment set, which in turn leads to better discriminability in image representation.
— *Geometric invariance* – the invariants w.r.t. geometric transformation, e.g., rotation, scaling, translation, and flipping, can be derived from the moment set, meaning better robustness in image representation.

Moments and moment invariants were introduced to the pattern recognition communities in 1962 by Hu [10]. Since then, after almost 60 years of research, numerous moment-based techniques have been developed for image representation with varying degrees of success. In 1998, Mukundan and Ramakrishnan [11] surveyed the main publications proposed until then and summarized the theoretical aspects of several classical moment functions. In 2006, Pawlak [12] gave a comprehensive survey on the reconstruction and calculation aspects of the moments with great emphasis to the accuracy/error analysis. In 2007, Shu et al. [13–15] provided a brief literature review for the mathematical definitions, invariants, and fast/accurate calculations of the classical moments, respectively. In 2009, Flusser et al. [8] presented a unique overview of moment-based pattern recognition methods with significant contribution to the theory of moment invariants. The substantial expansion [16] of this book includes more detailed analysis of the 3D object invariant representation. In 2011, Hoang [17] reviewed unit disk–based orthogonal moments in his doctoral dissertation, covering theoretical analysis, mathematical properties, and specific implementation. For most of the above reviews, state-of-the-art methods in the past 10 years are not covered. In 2014, Papakostas et al. [18] gave a global overview of the milestones in the 50 years research and highlighted all recent rising topics in this field. However, the theoretical basis for these latest research directions is rarely introduced. More recently, in 2019, Kaur et al. [19] provided a comparative review for many classical and new moments. Although this article covers almost all the main literatures, it still lacks the overall analysis of the current research progress in various directions. A common weakness of all the works is that there are almost no available software packages, restricting the further development of the community.

The significant contribution of this article is to give a systematic investigation for orthogonal moments–based image representation along with an open source implementation, which we believe would be a useful complement to [17–19]. For completeness, this article starts with some basic theories and classical methods in the area of orthogonal moments. Different from the mentioned reviews, we pay special attention to the motivation and successful experiences behind these traditional works. Furthermore, we organize a discussion for the recent advances of orthogonal moments on different research directions, including fast/accurate calculation, robustness/invariance optimization, definition extension, and application. Such overall theoretical analysis of the state-of-the-art research progress is mostly ignored in previous studies. In addition, we show the performance evaluation of widely used orthogonal moments in terms of moment calculation, image reconstruction, and pattern recognition. To embrace the concept of reproducible research, the corresponding software package is available online. In the end, several promising directions for future research are given along with some initial discussions/suggestions.

The rest of this article is organized as follows. In the following Section 2, we first give the basic idea of image moments and categorize existing methods into different categories. In Section 3, we further review in detail the unit disk–based orthogonal moments that are most relevant to image representation. Then, the recent advances of orthogonal moments within each research direction are reviewed and analyzed in Section 4. Furthermore, Section 5 reports the comparative results of state-of-the-art orthogonal moments along with an open source implementation. Due to space limitations, the detailed information of Section 5 is presented in the online Supplementary Material. Section 6 gives a conclusion and highlights some promising directions in this field.

2 OVERVIEW

Mathematically, the image moment is generally defined as the inner product $\langle f, V_{nm} \rangle$ of the image function f and the basis function V_{nm} of $(n+m)$ order on the domain D [8]:

$$\langle f, V_{nm} \rangle = \iint_D V_{nm}^*(x,y) f(x,y) dx dy, \tag{1}$$

where the asterisk $*$ denotes the complex conjugate. The direct geometric interpretation of image moment set $\langle f, V_{nm} \rangle$ is that it is the projection of f onto a subspace formed by a set of basis functions $\{V_{nm} : (n,m) \in \mathbb{Z}^2\}$ [17]. Because there is an infinite number of basis function sets, it is

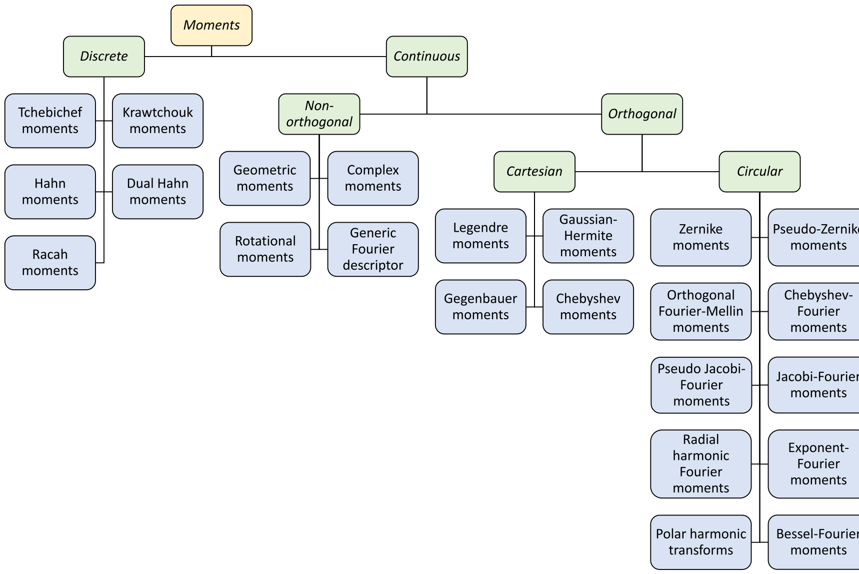


Fig. 2. A classification of image moments.

often necessary to manually design a set of special basis functions V_{nm} with beneficial properties in $\langle f, V_{nm} \rangle$ that meet the semantic requirements. According to the mathematical properties of basis functions, the family of image moments can be divided into different categories, as shown in Figure 2.

Firstly, depending on whether the basis functions satisfy *orthogonality*, the image moments can be classified into orthogonal moments and non-orthogonal moments. The orthogonality means any two different basis functions V_{nm} and $V_{n'm'}$ from the basis function set are uncorrelated or they are "perpendicular" in geometric terms, leading to no redundancy in the moment set. Mathematically, V_{nm} and $V_{n'm'}$ are orthogonal when the following condition is satisfied:

$$\langle V_{nm}, V_{n'm'} \rangle = \iint_D V_{nm}(x,y) V^*_{n'm'}(x,y) dxdy = \delta_{nn'} \delta_{mm'}, \tag{2}$$

where δ_{ij} is the Kronecker delta function defined as

$$\delta_{ij} = \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases}. \tag{3}$$

Some of the most popular non-orthogonal moments are geometric moments [8], rotational moments [20], complex moments [21], and generic Fourier descriptor [22]. Due to the non-orthogonality of the basis functions, high information redundancy exists in such moments. This further leads to difficulties in image reconstruction and poor discriminability/robustness in image representation. Therefore, it is a natural requirement to satisfy orthogonality when designing basis functions.

Secondly, according to whether the basis function is *continuous*, the image moments can be divided into continuous moments and discrete moments. In the case of two-dimensional (2D) images, the basis functions for continuous and discrete moments are generally defined in the 2D real-valued space and the 2D digital image space, i.e., the domains $D \in \mathbb{R}^2$ and $D \in \mathbb{Z}^2$, respectively. When it is necessary to calculate the continuous moments of a digital image, a suitable discretization approximation of the continuous integral is often introduced with corresponding computational errors. In Section 4.1, we will further describe the causes and solutions to such errors. On the contrary, discrete moments, such as Tchebichef moments [23], Krawtchouk moments [24], Hahn moments [25], dual Hahn moments [26], and Racah moments [27], do not involve any approximation errors. Thus, they are more suitable for the high-precision image processing tasks, e.g., image reconstruction, compression, and denoising.

Finally, depending on the *coordinate system* that defines the basis functions, image moments can be grouped into Cartesian moments and circular moments. In the case of continuous moments, the basis functions of Cartesian moments are defined in $D = \{(x, y) : x \in (-\infty, +\infty), y \in (-\infty, +\infty)\}$ or $D = \{(x, y) : x \in [-1, 1], y \in [-1, 1]\}$, while the domain for the circular moments is $D = \{(r, \theta) : r \in [0, +\infty), \theta \in [0, 2\pi)\}$ or $D = \{(r, \theta) : r \in [0, 1], \theta \in [0, 2\pi)\}$ (i.e., the unit disk). According to the proof of Bhatia and Wolf [28], the basis function V_{nm} will be *invariant in form* w.r.t. rotations of axes about the origin $(x, y) = (0, 0)$ if and only if, when expressed in polar coordinates (r, θ), it is of the form

$$V_{nm}(r\cos\theta, r\sin\theta) \equiv V_{nm}(r, \theta) = R_n(r)A_m(\theta), \tag{4}$$

with *angular basis function* $A_m(\theta) = \exp(\boldsymbol{j}m\theta)$ ($\boldsymbol{j} = \sqrt{-1}$) and *radial basis function* $R_n(r)$ could be of any form [17]. Let f^{rot} be the rotated version of the original image f. When V_{nm} conforms to the form of Equation (4), there must be a function $\mathcal{I}$ such that

$$\mathcal{I}\left(\left\{\langle f, V_{nm}\rangle\right\}\right) \equiv \mathcal{I}\left(\left\{\langle f^{\text{rot}}, V_{nm}\rangle\right\}\right), \tag{5}$$

i.e., satisfying the *rotation invariance*. Therefore, the Cartesian moments, such as Legendre moments [29], Gaussian-Hermite moments [30], Gegenbauer moments [31], Chebyshev moments [32], and the discrete moments listed above, have difficulties in achieving the rotation invariance. As for the calculation of circular moments, an appropriate coordinate transformation is often introduced since digital images are generally defined in a Cartesian coordinate system with corresponding computational errors. In Section 4.1, we will further describe the causes and solutions to such errors.

From the above theoretical analysis, it is clear that the circular orthogonal moments are generally better than other kinds of moments as far as image representation tasks are concerned. Therefore, great scientific interest has been given to the circular orthogonal moments, mainly the unit disk–based orthogonal moments. As the most relevant works, existing unit disk–based orthogonal moments will be reviewed in the next section.

3 CLASSICAL ORTHOGONAL MOMENTS

It can be checked from Equation (4) that the basis functions of unit disk–based orthogonal moments are separable, i.e., decomposed into the product of the radial basis functions and the angular basis functions. Therefore, Equation (2) can be rewritten as

$$\begin{aligned}\langle V_{nm}, V_{n'm'}\rangle &= \int_0^{2\pi}\int_0^1 R_n(r)A_m(\theta)R^*_{n'}(r)A^*_{m'}(\theta)rdrd\theta = \int_0^{2\pi} A_m(\theta)A^*_{m'}(\theta)d\theta \int_0^1 R_n(r)R^*_{n'}(r)rdr \\ &= 2\pi\delta_{mm'}\int_0^1 R_n(r)R^*_{n'}(r)rdr.\end{aligned} \tag{6}$$

Table 1. Definitions of Radial Basis Functions of Unit Disk–Based Orthogonal Moments

Method	Radial Basis Function
ZM	$R_{nm}^{(\mathrm{ZM})}(r) = \sqrt{\frac{n+1}{\pi}} \sum_{k=0}^{\frac{n-\vert m\vert}{2}} \frac{(-1)^k (n-k)! r^{n-2k}}{k!(\frac{n+\vert m\vert}{2}-k)!(\frac{n-\vert m\vert}{2}-k)!}$
PZM	$R_{nm}^{(\mathrm{PZM})}(r) = \sqrt{\frac{n+1}{\pi}} \sum_{k=0}^{n-\vert m\vert} \frac{(-1)^k (2n+1-k)! r^{n-k}}{k!(n+\vert m\vert+1-k)!(n-\vert m\vert-k)!}$
OFMM	$R_{n}^{(\mathrm{OFMM})}(r) = \sqrt{\frac{n+1}{\pi}} \sum_{k=0}^{n} \frac{(-1)^{n+k} (n+k+1)! r^{k}}{k!(n-k)!(k+1)!}$
CHFM	$R_{n}^{(\mathrm{CHFM})}(r) = \frac{2}{\pi} \left(\frac{1-r}{r}\right)^{\frac{1}{4}} \sum_{k=0}^{\lfloor \frac{n}{2} \rfloor} \frac{(-1)^{k} (n-k)! (4r-2)^{n-2k}}{k!(n-2k)!}$
PJFM	$R_{n}^{(\mathrm{PJFM})}(r) = \sqrt{\frac{(n+2)(r-r^2)}{\pi(n+3)(n+1)}} \sum_{k=0}^{n} \frac{(-1)^{n+k} (n+k+3)! r^{k}}{k!(n-k)!(k+2)!}$
JFM	$R_{n}^{(\mathrm{JFM})}(p,q,r) = \sqrt{\frac{r^{q-2}(1-r)^{p-q}(p+2n)\cdot\Gamma(q+n)\cdot n!}{2\pi\Gamma(p+n)\cdot\Gamma(p-q+n+1)}} \sum_{k=0}^{n} \frac{(-1)^{k}\Gamma(p+n+k) r^{k}}{k!(n-k)!\Gamma(q+k)}$
RHFM	$R_{n}^{(\mathrm{RHFM})}(r) = \begin{cases} \frac{1}{\sqrt{2\pi r}} & n=0 \\ \sqrt{\frac{1}{\pi r}} \sin(\pi(n+1)r) & n>0 \ \&\ n \text{ odd} \\ \sqrt{\frac{1}{\pi r}} \cos(\pi n r) & n>0 \ \&\ n \text{ even} \end{cases}$
EFM	$R_{n}^{(\mathrm{EFM})}(r) = \frac{1}{\sqrt{2\pi r}} \exp(\boldsymbol{j} 2n\pi r)$
PCET	$R_{n}^{(\mathrm{PCET})}(r) = \frac{1}{\sqrt{\pi}} \exp(\boldsymbol{j} 2n\pi r^2)$
PCT	$R_{n}^{(\mathrm{PCT})}(r) = \begin{cases} \frac{1}{\sqrt{\pi}} & n=0 \\ \sqrt{\frac{2}{\pi}} \cos(n\pi r^2) & n>0 \end{cases}$
PST	$R_{n}^{(\mathrm{PST})}(r) = \sqrt{\frac{2}{\pi}} \sin(n\pi r^2)$
BFM	$R_{n}^{(\mathrm{BFM})}(r) = \frac{1}{\sqrt{\pi} J_{v+1}(\lambda_n)} J_v(\lambda_n r),\ J_v(x) = \sum_{k=0}^{\infty} \frac{(-1)^k}{k!\Gamma(v+k+1)} \left(\frac{x}{2}\right)^{v+2k}$

Since $\langle V_{nm}, V_{n'm'} \rangle = \delta_{nn'}\delta_{mm'}$, the radial basis function $R_n(r)$ should satisfy the following weighted orthogonality condition:

$$\int_0^1 R_n(r) R_{n'}^*(r) r dr = \frac{1}{2\pi} \delta_{nn'}. \tag{7}$$

Equation (7) is a general requirement that must be considered when designing unit disk–based orthogonal moments, which ensures that the designed basis function set has the beneficial orthogonal properties. The angular basis functions $A_m(\theta)$ have a fixed form $\exp(\boldsymbol{j}m\theta)$ due to the proof of Bhatia and Wolf [28], which means that the difference in existing methods is only in the definition of the radial basis functions. In this regard, there are mainly three types of orthogonal functions used as the definition, including *Jacobi polynomials*, *harmonic functions*, and *eigenfunctions*. Next, we will briefly introduce the specific methods in these three groups and give their radial basis function definitions in a unified form, i.e., normalized version as Equation (7).

3.1 Jacobi Polynomials

In this group, the famous methods mainly include ZM [29], **Pseudo-Zernike Moments (PZM)** [20], **Orthogonal Fourier-Mellin Moments (OFMM)** [33], **Chebyshev-Fourier Moments (CHFM)** [34], **Pseudo Jacobi-Fourier Moments (PJFM)** [35], and **Jacobi-Fourier Moments (JFMs)** [36]. Their radial basis function definitions are summarized in Table 1, which directly satisfy the orthogonality condition in Equation (7).

It is worth mentioning that the radial basis function of JFM is constructed directly from the original Jacobi polynomials [28], while the radial basis functions of ZM, PZM, OFMM, CHFM, and PJFM are all special cases of the Jacobi polynomials. Thus, JFM is, in fact, a generic expression of the above famous methods. By properly setting the values of the parameters p and q, $R_{nm}^{(\text{OFMM})}(r)$, $R_{nm}^{(\text{CHFM})}(r)$, and $R_{nm}^{(\text{PJFM})}(r)$ can be directly obtained from $R_{nm}^{(\text{JFM})}(p,q,r)$. The relationship between $R_{nm}^{(\text{JFM})}(p,q,r)$ and $R_{nm}^{(\text{ZM})}(r)/R_{nm}^{(\text{PZM})}(r)$ is more complicated, and we refer readers to the work of Hoang Tabbone [37] for more details. Here, the parameter setting is listed below:

- ZM – $p = |m| + 1$ and $q = |m| + 1$;
- PZM – $p = 2|m| + 2$ and $q = 2|m| + 2$;
- OFMM – $p = 2$ and $q = 2$;
- CHFM – $p = 2$ and $q = 1.5$;
- PJFM – $p = 4$ and $q = 3$.

The unit disk–based orthogonal moments using Jacobi polynomials, especially the landmark ZM, have a long history in optical physics, digital image processing, and pattern recognition. As one can note from the formulas listed above, however, the definitions of these radial basis functions rely on factorial/gamma terms and summations, which leads to high computational complexity. In addition, the factorial and gamma functions tend to cause the calculation errors, mainly *numerical instability*. In Sections 4.1 and 4.2, we will further describe the causes and solutions to the above issues.

3.2 Harmonic Functions

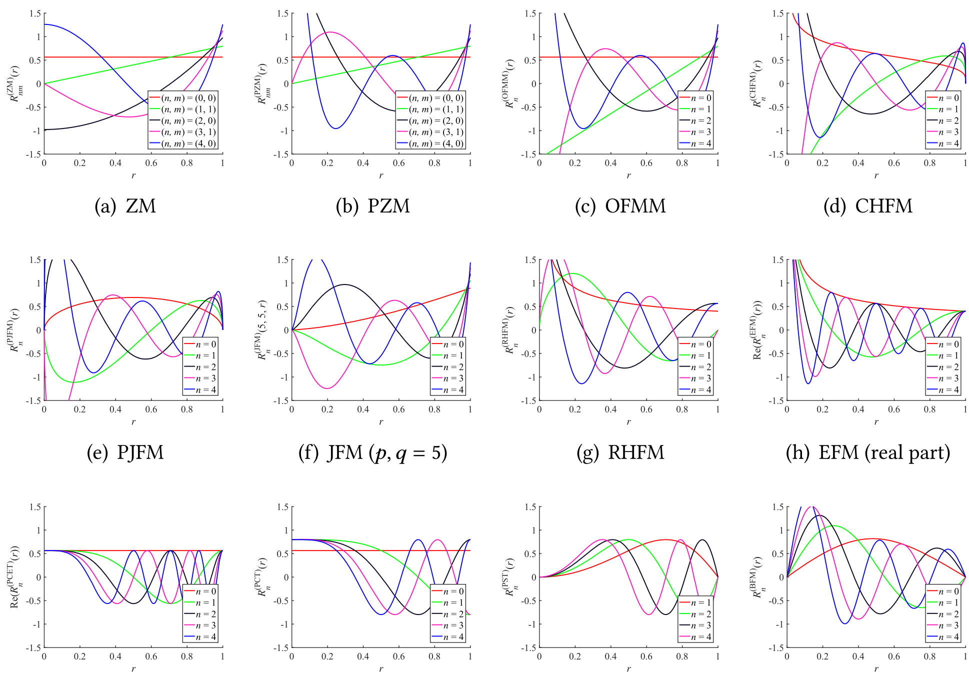
In this group, the famous methods mainly include **Radial Harmonic Fourier Moments** (**RHFMs**) [38], **Exponent-Fourier Moments** (**EFMs**) [39], and PHT [40]. Here, PHT consists of three different transformations: **Polar Complex Exponential Transform** (**PCET**), **Polar Cosine Transform** (**PCT**), and **Polar Sine Transform** (**PST**). Their radial basis function definitions are summarized in Table 1, which directly satisfy the orthogonality condition in Equation (7).

It can be seen that the radial basis functions of RHFMs, EFMs, and PHTs are all based on the harmonic functions commonly used in Fourier analysis, i.e., complex exponential functions $\{\exp(\boldsymbol{j}2n\pi r) : n \in \mathbb{Z}\}$ and trigonometric functions $\{1, \cos(2n\pi r), \sin(2n\pi r) : n \in \mathbb{Z}^+\}$. Therefore, the definitions of the above methods are closely related, and the radial basis functions can be transformed into each other via Euler's formula $\exp(\boldsymbol{j}\alpha) = \cos(\alpha) + \boldsymbol{j}\sin(\alpha)$ and variable substitution $dr = \frac{1}{2}dr^2$. More details on this will be given in Section 4.4. As for the calculation, compared to Jacobi polynomials, the orthogonal moments using only harmonic functions do not involve any complicated factorial/gamma terms and long summations, meaning better time complexity and numerical stability.

3.3 Eigenfunctions

At current stage, there are still relatively few orthogonal moments based on eigenfunctions, and the most representative one is **Bessel-Fourier Moments** (**BFMs**) [41]. The radial basis function definition of BFM is included in Table 1, which directly satisfies the orthogonality condition in Equation (7).

It is observed that the radial basis function relies on infinite series and factorial/gamma terms, and also requires a root-finding algorithm for Bessel functions of the first kind $J_v(x)$ to determine



(a) ZM (b) PZM (c) OFMM (d) CHFM

(e) PJFM (f) JFM ($p, q = 5$) (g) RHFM (h) EFM (real part)

(i) PCET (real part) (j) PCT (k) PST (l) BFM

Fig. 3. Illustrations of radial basis functions of unit disk–based orthogonal moments.

the n-th zero λ_n. Therefore, compared with the Jacobi polynomials and harmonic functions, the orthogonal moments based on eigenfunctions have significantly higher complexity in theory.

3.4 Summary and Discussion

For a better perception, the illustrations of radial basis functions of unit disk–based orthogonal moments are summarized in Figure 3.

Furthermore, we reveal the mathematical properties of these methods in Table 2, including radial basis functions' parameter, computational complexity, numerical stability, number of zeros, and distribution of zeros. The complexity of basis functions (considering only the definition-style computation) is graded as low, high, and very high based on whether the definition involves factorial/gamma terms, summation/series operations, and root-finding processes. Numerical stability is graded as poor, medium, and good depending on whether the function includes factorial/gamma terms and very high absolute values (mainly unbounded). The number of zeros of radial kernels is related to the ability of moments for capturing image high-frequency information [17, 40]. Besides the quantity, radial kernel's other key attribute is the distribution of zeros, because it is related to the description emphasis of the moments in image plane [17, 40]. When the essential discriminative information is distributed uniformly in the spatial domain, unfair emphasis of the extracted moments on certain regions has been shown to have a negative impact on the discrimination quality, termed as *information suppression problem* [21].

It can be seen in Table 2 that, among these unit disk–based orthogonal moments, almost no method has both low complexity, good stability, large number, and unbiased zeros. This observation strongly motivates the design of the improvement strategies that address the above common shortcomings, which will be discussed in Section 4.

Table 2. Mathematical Properties of Radial Basis Functions of Unit Disk-Based Orthogonal Moments

Method	Parameter	Complexity	Stability	Number of Zeros	Distribution of Zeros
ZM	$n \in \mathbb{N}, m \in \mathbb{Z}, \|m\| \leq n,$ $n - \|m\| = \text{even}$	high	poor	$(n - \|m\|)/2$	biased
PZM	$n \in \mathbb{N}, m \in \mathbb{Z}, \|m\| \leq n$	high	poor	$n - \|m\|$	biased
OFMM	$n \in \mathbb{N}, m \in \mathbb{Z}$	high	poor	n	basically uniform
CHFM	$n \in \mathbb{N}, m \in \mathbb{Z}$	high	poor	n	basically uniform
PJFM	$n \in \mathbb{N}, m \in \mathbb{Z}$	high	poor	n	basically uniform
JFM	$n \in \mathbb{N}, m \in \mathbb{Z}, p, q \in \mathbb{R},$ $p - q > -1, q > 0$	high	poor	n	basically uniform
RHFM	$n \in \mathbb{N}, m \in \mathbb{Z}$	low	medium	n	uniform
EFM	$n \in \mathbb{Z}, m \in \mathbb{Z}$	low	medium	$2n$	uniform
PCET	$n \in \mathbb{Z}, m \in \mathbb{Z}$	low	good	$2n$	biased
PCT	$n \in \mathbb{N}, m \in \mathbb{Z}$	low	good	n	biased
PST	$n \in \mathbb{N}^+, m \in \mathbb{Z}$	low	good	$n - 1$	biased
BFM	$n \in \mathbb{N}, m \in \mathbb{Z}, v \in \mathbb{R},$ $\lambda_n = n\text{-th zero of } J_v(x)$	very high	medium	n	basically uniform

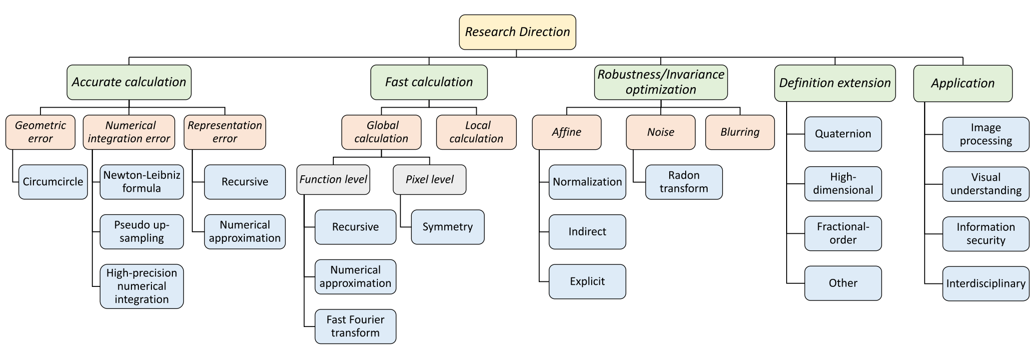


Fig. 4. The research directions of image moments.

4 RESEARCH DIRECTIONS AND RECENT ADVANCES

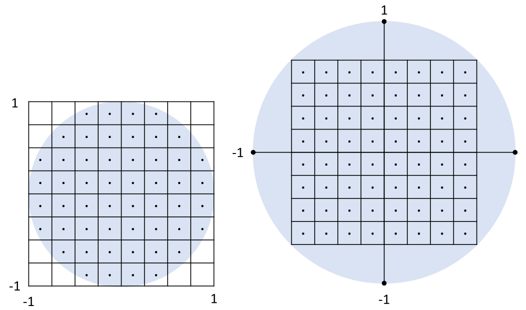
In addition to defining new image moments, existing work mainly focuses on optimizing the classical moments listed in Sections 2 and 3. As shown in Figure 4, the current research directions generally include accurate calculation, fast calculation, robustness/invariance optimization, definition extension, and application. This section will introduce the above directions along with the recent advances.

4.1 Accurate Calculation

As the mathematical background of Sections 4.1 and 4.2, the general procedure for calculating image moments is first given.

Going back to Equation (1), for computing the image moments $\langle f, V_{nm} \rangle$ of a digital image defined in the discrete Cartesian grid $\{f(i,j) : (i,j) \in [1, 2, \ldots, N]^2\}$, it is first necessary to unify the domains of $f(i,j)$ and $V_{nm}(x,y)$, i.e., to design a mapping between the two:

$$(i,j) \rightarrow (x_i, y_j), \tag{8}$$



(a) Incircle (b) Circumcircle

Fig. 5. The coordinate mapping between the square region of digital image and the unit disk.

where a pixel region $[i - \frac{\Delta i}{2}, i + \frac{\Delta i}{2}] \times [j - \frac{\Delta j}{2}, j + \frac{\Delta j}{2}]$ centered at (i, j) is mapped into a region $[x_i - \frac{\Delta x_i}{2}, x_i + \frac{\Delta x_i}{2}] \times [y_j - \frac{\Delta y_j}{2}, y_j + \frac{\Delta y_j}{2}]$ centered at (x_i, y_j). By the coordinate mapping $(i, j) \rightarrow (x_i, y_j)$, Equation (1) can be written down in discrete form as follows:

$$\langle f, V_{nm} \rangle = \sum_{(x_i, y_j) \in D} h_{nm}(x_i, y_j) f(i, j), \tag{9}$$

where $h_{nm}(x_i, y_j)$ is the integral value of the basis function V_{nm} over the mapped pixel region $[x_i - \frac{\Delta x_i}{2}, x_i + \frac{\Delta x_i}{2}] \times [y_j - \frac{\Delta y_j}{2}, y_j + \frac{\Delta y_j}{2}]$, defined as

$$h_{nm}(x_i, y_j) = \int_{x_i - \frac{\Delta x_i}{2}}^{x_i + \frac{\Delta x_i}{2}} \int_{y_j - \frac{\Delta y_j}{2}}^{y_j + \frac{\Delta y_j}{2}} V_{nm}^*(x, y) dx dy. \tag{10}$$

In general, the calculation of unit disk–based orthogonal moments suffers from *geometric error*, *numerical integration error*, and *representation error* (mainly numerical instability). These errors will severely restrict the quality of image representation, especially when high-order moments are required to better describe the image. Hence, the accurate computation strategies of moments are vital for the applicability.

4.1.1 Geometric Error. Geometric error may occur when mapping the image domain into the basis function domain, i.e., Equation (8). Such errors are common in the calculation of the unit disk–based orthogonal moments, because digital images are generally defined over a square region in the Cartesian coordinate system, rather than the unit disk.

As for mapping between the square region and the unit disk, there are naturally, as shown in Figure 5, the *incircle* mapping [42]:

$$\begin{cases} i \rightarrow x_i = \frac{2i-N}{N} = i\Delta x_i - 1 \\ j \rightarrow y_j = \frac{2j-N}{N} = j\Delta y_j - 1 \end{cases}, \tag{11}$$

and the *circumcircle* mapping [43]:

$$\begin{cases} i \rightarrow x_i = \frac{2i-N}{\sqrt{2}N} = i\Delta x_i - \frac{\sqrt{2}}{2} \\ j \rightarrow y_j = \frac{2j-N}{\sqrt{2}N} = j\Delta y_j - \frac{\sqrt{2}}{2} \end{cases}. \tag{12}$$

When the incircle mapping is used, there exists (i, j) such that $(x_i, y_j) \notin D$, i.e., the pixels that are mapped to the outside of the unit disk are not counted, and there exists (i, j) such that

$[x_i - \frac{\Delta x_i}{2}, x_i + \frac{\Delta x_i}{2}] \times [y_j - \frac{\Delta y_j}{2}, y_j + \frac{\Delta y_j}{2}] - D \neq \emptyset$, i.e., some mapped pixel regions partially intersect the unit disk. In both cases, geometric errors occur.

As for the circumcircle mapping, it is able to completely avoid such two cases, so there will be no geometric errors. However, this comes at the cost of representation capability [17], as the mapped regions containing image information occupy only $\frac{2}{\pi}$ of the entire unit disk.

4.1.2 Numerical Integration Error. Numerical integration error may occur when calculating the integral of the continuous basis functions, i.e., Equation (10). Here, for discrete moments, the basis functions V_{nm} are generally constant on the interval $[x_i - \frac{\Delta x_i}{2}, x_i + \frac{\Delta x_i}{2}] \times [y_j - \frac{\Delta y_j}{2}, y_j + \frac{\Delta y_j}{2}]$. Thus, it is easy to check that $h_{nm}(x_i, y_j) = V_{nm}^*(x_i, y_j)\Delta x_i \Delta y_j$, meaning discrete moments typically do not involve numerical integration errors [23]. As for continuous moments, since the basis functions V_{nm} are continuous over the interval $[x_i - \frac{\Delta x_i}{2}, x_i + \frac{\Delta x_i}{2}] \times [y_j - \frac{\Delta y_j}{2}, y_j + \frac{\Delta y_j}{2}]$, solving $h_{nm}(x_i, y_j)$ requires some integration tricks.

As the simplest case, there is an analytical solution to $h_{nm}(x_i, y_j)$, i.e., Equation (10) can be solved directly by the *Newton-Leibniz formula* [44, 45]. Such calculations also do not involve numerical integration errors.

As the more general case, considering the complication of the definition of many basis functions V_{nm}, it is often difficult to determine the analytical solution and some approximate algorithms are needed for achieving the numerical solution of $h_{nm}(x_i, y_j)$. The most commonly used approximation algorithm is ***Zero-Order Approximation*** (**ZOA**) [42], which imitates the calculation of discrete moments, as follows:

$$h_{nm}(x_i, y_j) \simeq V_{nm}^*(x_i, y_j)\Delta x_i \Delta y_j. \tag{13}$$

Generally, the accuracy of the numerical integration method is inversely proportional to the interval area $\Delta x_i \Delta y_j$. Therefore, by further dividing a single pixel region into multiple smaller integration intervals (e.g., 3×3 sub-intervals), higher accuracy can be easily obtained. This strategy is often called *pseudo up-sampling* [46, 47]. When the ZOA is used in these sub-intervals, the integral can be expressed as

$$h_{nm}(x_i, y_j) \simeq \sum_{(a,b)} V_{nm}^*(u_a, v_b)\Delta u_a \Delta v_b, \tag{14}$$

where (u_a, v_b) is the sampling point with $u_a \in [x_i - \frac{\Delta x_i}{2}, x_i + \frac{\Delta x_i}{2}]$ and $v_b \in [y_j - \frac{\Delta y_j}{2}, y_j + \frac{\Delta y_j}{2}]$. In addition to this simple approach, other more complex *high-precision numerical integration* strategies [48–50], such as the *Gaussian quadrature rule* and *Simpson's rule*, can also be used in the computation of Equation (10). Their general definition is

$$h_{nm}(x_i, y_j) \simeq \sum_{(a,b)} w_{ab} V_{nm}^*(u_a, v_b)\Delta x_i \Delta y_j, \tag{15}$$

where w_{ab} is the weight corresponding to the sampling point (u_a, v_b).

It is worth noting that all of the above discussion in Section 4.1 relies on the Cartesian coordinate system, which we call Cartesian-based calculation method. Correspondingly, there also exists a calculation method based on a polar coordinate system for unit disk–based orthogonal moments, often referred to as *polar pixel tiling* [51, 52].

It first resamples the digital image $f(i, j)$ to a discrete polar grid (r_u, θ_{uv}), as shown in Figure 6, and then performs calculation in a manner similar to Equation (9) and Equation (10):

$$\langle f, V_{nm} \rangle = \sum_{(r_u, \theta_{uv}) \in D} h_{nm}(r_u, \theta_{uv}) f(r_u, \theta_{uv}), \tag{16}$$

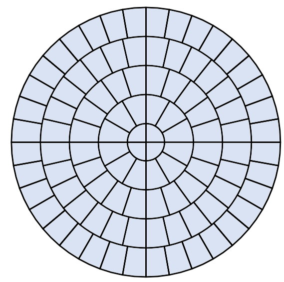


Fig. 6. The coordinate mapping in polar pixel tiling scheme.

with

$$h_{nm}(r_u, \theta_{uv}) = \int_{r_u}^{r_{u+1}} \int_{\theta_{uv}}^{\theta_{u(v+1)}} V_{nm}^*(r, \theta) r dr d\theta = \int_{r_u}^{r_{u+1}} R_n(r) r dr \int_{\theta_{uv}}^{\theta_{u(v+1)}} \exp(-\boldsymbol{j} m\theta) d\theta. \tag{17}$$

Note that Equation (17) has a distinct advantage, i.e., the computation of $h_{nm}(r_u, \theta_{uv})$ can be separated into two independent parts [53]: (1) $\int_{r_u}^{r_{u+1}} R_n(r) r dr$ can be approximately integrated by pseudo up-sampling and the Gaussian quadrature rule; (2) $\int_{\theta_{uv}}^{\theta_{u(v+1)}} \exp(-\boldsymbol{j} m\theta) d\theta$ can be exactly integrated by the Newton-Leibniz formula, as follows:

$$\int_{\theta_{uv}}^{\theta_{u(v+1)}} \exp(-\boldsymbol{j} m\theta) d\theta = \begin{cases} \frac{\boldsymbol{j}[\exp(-\boldsymbol{j} m\theta_{u(v+1)}) - \exp(-\boldsymbol{j} m\theta_{uv})]}{m} & m \neq 0 \\ \theta_{u(v+1)} - \theta_{uv} & m = 0 \end{cases}. \tag{18}$$

In addition to the above advantage, the polar pixel tiling has similar properties to the Cartesian-based calculation method in terms of geometric error and numerical integration error, and will not be repeated here.

4.1.3 Representation Error. Representation error is caused by the finite precision of the numerical computing systems, which occurs in all processes of the calculation, i.e., Equation (8), Equation (9), and Equation (10). The representation error can be further divided into *overflow error*, *underflow error*, and *roundoff error* [17]. The common numerical instability is mainly attributable to roundoff error and overflow error.

The basis functions based on Jacobi polynomials contain factorial/gamma terms. When the order is large, the actual values of these coefficients may exceed the representation range of the numerical computing system, resulting in roundoff error or even overflow error. To avoid direct calculation of the factorial/gamma terms, *recursive* strategies [54–56] and *numerical approximation* algorithms [57, 58] are often used. The recursive method relies on the recursive relationship of the basis functions derived from $a! = a \cdot (a-1)!$ and $\Gamma(a) = a \cdot \Gamma(a-1)$, using several low-order basis functions to directly derive the high-order ones. This process does not involve the factorial/gamma of large number. In another path, the numerical approximation method achieves factorial-free calculations by resorting to a suitable approximation for factorials such as *Stirling's formula*.

In addition to factorial/gamma terms, the basis functions of some orthogonal moments have very high (or even infinite) absolute values at certain points, which may also exceed the representation

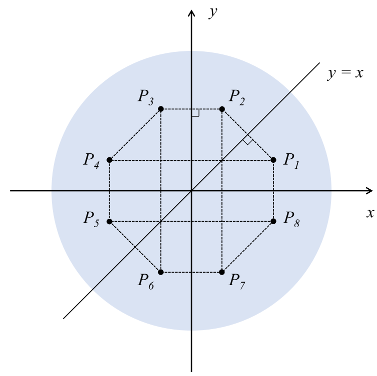


Fig. 7. Illustrations of symmetrical points in the unit disk.

range of the numerical computing system. For example, although the definitions of RHFM and EFM do not involve any factorial/gamma terms, their basis functions are infinite at the origin (see also Figure 3), which will cause roundoff error or even overflow error.

4.1.4 Recent Advance of Accurate Calculation. In the field of accurate calculation for unit disk–based orthogonal moments, Xin et al.'s polar pixel tiling [51] may be the most influential and representative work. Recent papers typically combine polar pixel tiling with other techniques, including circumcircle mapping, pseudo up-sampling, high-precision numerical integration, Newton-Leibniz formula, and recursive strategy, to provide the state-of-the-art performance.

In this way, Sáez-Landete [59] integrated the work of Camacho-Bello et al. [53] and Upneja and Singh [56], giving an accurate calculation algorithm for the JFM. Hence, it also applies to other Jacobi polynomial-based moments such as OFMM and CHFM. Due to the polar pixel tiling, the complicated radial parts can be evaluated using Gaussian quadrature, pseudo up-sampling, and recursive relationship, while the angular parts can be exactly evaluated by the Newton-Leibniz formula. A somewhat similar method was proposed by Hosny and Darwish [60] for the harmonic function–based PHT. The main difference is that, due to the simple definition of the radial basis function, the radial parts can also be integrated exactly by the Newton-Leibniz formula just like the angular parts.

These two recent works provide near-perfect accuracy for calculating Jacobi polynomial-based moments and harmonic function-based moments, respectively. To be critical, the only flaw may be the error introduced in the process of image resampling to discrete polar grid [51]. If the nonlinear interpolation methods are used, such as bicubic method, the interpolation error is negligible for image representation tasks. On the other hand, if a given task is sensitive to such errors, other more complex mathematical tools, such as pseudo-polar [61], are instructive for the design of accurate calculations [62].

4.2 Fast Calculation

In the implementation of orthogonal moments, the overall complexity may become excessively high when (1) a large number of moments is needed, (2) the image has high resolution, (3) many images need to be processed, or (4) a high-precision computation is required. Since these requirements are common in practical applications, the fast calculation of orthogonal moments is strongly demanded. Depending on the application scenario, optimization efforts for computational speed can be divided into two categories: *global calculation* and *local calculation*.

4.2.1 Global Calculation. Global calculation is to calculate the image moments for the entire image. In this scenario, the number of moments and the resolution of the image are the main factors that affect the time complexity. Taking the simplest ZOA-based direct calculation as an example, the time cost of an n-order image moment comes from the following:

— For Equation (10), it is necessary to evaluate the values of the basis function V_{nm} at $N \times N$ sampling points, i.e., $\{V^*_{nm}(x_i, y_j) : (i, j) \in [1, 2, \ldots, N]^2\}$. The time complexity is related to the definition of the basis functions. For example, the basis functions using Jacobi polynomials require $O(nN^2)$ additions for the summation. In contrast, the basis functions using harmonic functions do not involve such additions.
— For Equation (9), it is necessary to calculate the inner product of the basis function V_{nm} and the digital image f at $N \times N$ sampling points, i.e., $\langle f(i, j), V^*_{nm}(x_i, y_j)\Delta x_i \Delta y_j\rangle$. The time complexity is $O(N^2)$ multiplications and $O(N^2)$ additions.

Note that the calculations listed above are only required for one moment. Let K be some integer constant; if all the moments of orders in set $\{(n, m) : |n|, |m| \leq K\}$ are computed, the total complexity will increase by a factor of $O(K^2)$. Moreover, the computational cost of Equation (10) may rise sharply if the high-precision numerical integration strategy, such as Gaussian quadrature, is adopted. To reduce the complexity of Equation (9) and Equation (10), the existing methods are designed at the *function level* and the *pixel level*.

— Function level: For the orthogonal moments based on Jacobi polynomials, if the recursive method and numerical approximation (see also Section 4.1.3) are used to evaluate the basis functions, the number of additions (from summation) and multiplications (from factorial) in the calculation of Equation (10) can be reduced [54–58]. For example, the recursive calculation of polynomial $R_n(\sqrt{x_i^2 + y_j^2})$ requires $O(N^2)$ additions, less than the $O(nN^2)$ in direct computation. If K polynomials $R_n(\sqrt{x_i^2 + y_j^2})$ of orders $\{0, 1, \ldots, K\}$ are required, the direct method involves $O(K^2N^2)$ additions, while the recursive scheme only requires $O(KN^2)$ additions. For the orthogonal moments based on harmonic functions, in addition to adopting the similar recursive form [63–66], a more effective strategy is to make use of its inherent relationship with Fourier transform [67–71]. Once the explicit relationship between the two is determined, the ***Fast Fourier Transform*** (**FFT**) algorithm can be used to calculate the moments. Note that FFT has the ability to reduce the number of the most time-consuming multiplications in Equation (9) from $O(N^2)$ to $O(N \log N)$.
— Pixel level: The most common strategy in this path is to simplify the calculation by exploring the *symmetry* and *anti-symmetry* of the basis functions in the domain [72, 73]. More specifically, the basis function values at all $N \times N$ sampling points can be completely derived by the basis function values at a few special sampling points, thus reducing the complexity of Equation (10) and Equation (9). In Figure 7, we give an illustration of the symmetrical points in the unit disk. For a point P_1, there are seven symmetrical points $\{P_2, P_3, P_4, P_5, P_6, P_7, P_8\}$ w.r.t. coordinate axes, origin, and $y = x$. Their Cartesian coordinates and polar coordinates are listed in Table 3. It can be seen that all these points have the same radial coordinate and related angular coordinate. Based on the mathematical properties of the complex exponential function, i.e., the trigonometric identities, such correlation of coordinates will be directly converted to the correlation of basis function values. Hence, this observation can lead to a reduction in computational complexity of Equation (10) and Equation (9) by approximately 1/8. Considering that all unit disk–based orthogonal moments are based on angular basis functions using complex exponential functions, the above symmetry

Table 3. Cartesian Coordinates and Polar Coordinates of the Symmetrical Points

Symmetrical Point	Symmetrical Axis	Cartesian Coordinates	Polar Coordinates
P_1		(x, y)	(r, θ)
P_2	$y = x$	(y, x)	$(r, \frac{\pi}{2} - \theta)$
P_3	$y = x$, y-axis	$(-y, x)$	$(r, \frac{\pi}{2} + \theta)$
P_4	y-axis	$(-x, y)$	$(r, \pi - \theta)$
P_5	origin	$(-x, -y)$	$(r, \pi + \theta)$
P_6	origin, $y = x$	$(-y, -x)$	$(r, \frac{3\pi}{2} - \theta)$
P_7	$y = x$, x-axis	$(y, -x)$	$(r, \frac{3\pi}{2} + \theta)$
P_8	x-axis	$(x, -y)$	$(r, 2\pi - \theta)$

and anti-symmetry properties are maintained in all these methods. As a result, this fast calculation strategy is generic and easy to use in combination with other fast algorithms.

4.2.2 Local Calculation. Local calculation is to calculate the image moments for a part of the image such as dense image blocks or interest regions of keypoints. In this application scenario, the resolutions of such image patches are generally small, which can be described well using few moments. As the main reason for pushing up the time complexity, the number of such patches is typically very large, e.g., this number for dense blocks can be of order 10^5 or even 10^6. Therefore, speed optimization methods for global computation often do not solve the local calculation problem well. At present, the problem of calculating local image moments has not been fully discussed. The most promising path is to find and compress redundant operations in the for-loop like processing of image patches. In this regard, *useful properties of basis functions* (such as shift properties in [74, 75]) and *special data structures* (such as complex-valued integral images in [76]) have been explored to improve efficiency.

4.2.3 Recent Advance of Fast Calculation. For the fast global calculation of Jacobi polynomial-based orthogonal moments, state-of-the-art methods usually use recursion and symmetry in combination. A recent work in this way was proposed by Upneja and Singh [56] for JFM (see also [53, 59]); it requires $\mathcal{O}(\frac{1}{8}(KN^2 + K^2N^2))$ additions and $\mathcal{O}(\frac{1}{8}K^2N^2)$ multiplications for all the moments of order in set $\{(n, m) : |n|, |m| \leq K\}$. Considering that multiplication is usually much more expensive than addition in modern computing systems, so the total elapsed time is mainly determined by the $\mathcal{O}(\frac{1}{8}K^2N^2) = \mathcal{O}(K^2N^2)$ multiplications. In this case, it is clear that when higher image resolution (i.e., N increases) or more moments (i.e., K increases) are required, the computational cost still increases in a *quadratic* manner. This is mainly due to the complicacy of the Jacobi polynomials, in other words, which leads to the lack of fast implementation with a complexity below the quadratic time.

For the fast global calculation of harmonic function–based orthogonal moments, the state-of-the-art performance offered by the FFT-based fast implementation. The earliest idea in this way can be traced back to 2014, when Ping et al. [70] introduced the FFT by the relationship between EFM and Fourier transform. Subsequently, similar ideas have been used for RHFM [68] and PCET [69]. An important work was recently proposed by Yang et al. [67], which generalized such FFT-based techniques to a generic version of harmonic function–based orthogonal moments (will be seen in Section 4.4.3). This generic fast implementation, similar to its previous special forms, exhibits the multiplicative complexity of $\mathcal{O}(M^2 \log M) = \mathcal{O}(N^2 \log N)$, where M is a sampling parameter and $M \propto N$. Note that, surprisingly, K is *no role* in the multiplicative complexity. This property means that when K is slightly higher, so that K^2 is greater than $\log N$, the elapsed time of this generic method will be significantly lower than the strategy based on recursion and symmetry such as

[67]. In addition to the efficiency gain, this method has also been proven (both analytically and experimentally) to avoid numerical instability, while providing quite high calculation accuracy when M is large.

For the fast local calculation, the recent work of Bera et al. [76] makes a crucial contribution. They reduced the ZM calculation of dense image blocks to *constant* time, i.e., $\mathcal{O}(1)$, by introducing an elegant data structure: complex-valued integral image. The further speed improvements can be achieved with the help of another structure: lookup table. While the use of integral image in fast implementation is not rare, e.g., SURF uses it to speed up SIFT [77], this is still the very first time such techniques have been used for orthogonal moments. Through the complicated derivation, we found that two mathematical tools played an important role in rewriting the definition of ZM toward integral images, namely, the complex-plane representation of radial and angular coordinates and the binomial expansion. The above structures and tools introduced by Bera et al. may provide important insights to researchers in the field. Unfortunately, as discussed by the authors, this fast algorithm seems to be applicable only for ZM and not for other moments with better representation power (i.e., no information suppression) such as OFMM. If the similar constant-time implementation could be developed for all the unit disk–based orthogonal moments, this will greatly promote their application in real-time tasks.

4.3 Robustness/Invariance Optimization

As the mathematical background of Section 4.3, the general definitions of invariance, robustness, and discriminability in image representation are first given.

If there exists a function $\mathcal{R}$ such that the original image f and its degraded version $\mathcal{D}(f)$, where $\mathcal{D}$ is the degradation operator, satisfy

$$\mathcal{R}(f) \equiv \mathcal{R}(\mathcal{D}(f)), \tag{19}$$

for any f, the representation $\mathcal{R}$ is said to be *invariant* to the degradation $\mathcal{D}$.

We consider function $\mathcal{L} : \mathbf{X} \times \mathbf{X} \to [0, +\infty)$ to be a *distance* measure on a set $\mathbf{X}$, where for all $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbf{X}$, the following three axioms are satisfied:

— *identity of indiscernibles* – $\mathcal{L}(\mathbf{x}, \mathbf{y}) = 0 \Leftrightarrow \mathbf{x} = \mathbf{y}$;
— *symmetry* – $\mathcal{L}(\mathbf{x}, \mathbf{y}) = \mathcal{L}(\mathbf{y}, \mathbf{x})$;
— *subadditivity* or *triangle inequality* – $\mathcal{L}(\mathbf{x}, \mathbf{y}) \leq \mathcal{L}(\mathbf{x}, \mathbf{z}) + \mathcal{L}(\mathbf{z}, \mathbf{y})$.

Given a distance function $\mathcal{L}$, *robustness* requires that the *intra-class distance* of the representation $\mathcal{R}$,

$$\mathcal{L}(\mathcal{R}(f), \mathcal{R}(\mathcal{D}(f))), \tag{20}$$

should be sufficiently small. Conversely, assuming that image g is semantically different from the image f, *discriminability* requires that the *inter-class distance* of the representation $\mathcal{R}$,

$$\mathcal{L}(\mathcal{R}(f), \mathcal{R}(g)), \tag{21}$$

should be sufficiently large. It can be seen that the invariance implies the perfect robustness, i.e., $\mathcal{L}(\mathcal{R}(f), \mathcal{R}(\mathcal{D}(f))) = 0$ holds if and only if Equation (19) holds due to the first axiom of $\mathcal{L}$. Here, the moment-based image representation $\mathcal{I}$ is a special form of the generic representation $\mathcal{R}$, which depends on the given basis function V_{nm} and can be written as

$$\mathcal{R}(f) = \mathcal{I}(\{\langle f, V_{nm} \rangle\}). \tag{22}$$

It is almost impossible to devise a representation that maintains well invariance or robustness to all kinds of degradations. More precisely, the only possibility corresponds to a representation without any discriminability [8]. Thus, in practice, the design of representation $\mathcal{R}$ generally relies on certain assumptions about the degradation $\mathcal{D}$. In general, invariance/robustness optimization methods are proposed against three types of attacks: *affine transformation*, *noise*, and *blurring*.

4.3.1 Affine Transformation. It is a transformation of image-space coordinates, that is, the degradation operator $\mathcal{D}$ is a mapping from the pixel coordinates (x, y) to the new coordinates (x', y'), $\mathcal{D} : (x, y) \to (x', y')$, and is defined as

$$\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \begin{pmatrix} a_{00} & a_{01} & t_x \\ a_{10} & a_{11} & t_y \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}, \tag{23}$$

where the affine parameters $(a_{00}, a_{01}, a_{10}, a_{11}, t_x, t_y)$ can encode rotation, scaling, translation, shearing, flipping, and all linear combinations of the above transformations. In the moment-based image representation, there are three methods to achieve invariance to the affine transformation $\mathcal{D}$: *normalization*, *indirect*, and *explicit* methods [8, 16, 78].

- Normalization method [79, 80]: In this approach, the degraded input image $\mathcal{D}(f)$ is converted to some *reference* form f^{ref}, so the representation of this normalized image, $\mathcal{R}(f^{\text{ref}})$, will be invariant to the affine transformation $\mathcal{D}$. Obviously, the invariance of this approach comes from the independent correction technique. More specifically, the normalization can be considered as a function $\mathcal{N}$, which must hold $\mathcal{N}(\mathcal{D}(f)) = \mathcal{N}(f) = f^{\text{ref}}$ for any image f and any admissible affine transformation $\mathcal{D}$. Note that the geometric transformation does not need to be actually implemented on the digital image; this normalization process is conceptual [8]. As can be seen, the normalization results will fundamentally affect the invariance of the representation. Over the years, the design of such a normalization approach has been pursued by many researchers in different tasks such as geometric correction [81, 82] and image registration [83, 84]. One of the simplest methods is based on geometric moments, which obtain translation, scale, and rotation invariance by evaluating/eliminating the *centroid*, *scaling factor*, and *principal axis* of the image, all based on low-order geometric moments. For more details on this, we strongly encourage readers to see the books by Flusser et al. [8, 16], where they provide the complete analysis and definition. For other types of normalization methods, readers can refer to the paper by Zitova and Flusser [84].
- Indirect method [85, 86]: It is well known that geometric moments are easy to derive explicit invariants for affine transformation $\mathcal{D}$, as shown in [8, 16, 87]. Therefore, this approach expresses orthogonal moment invariants as a linear combination of the geometric moment invariants, based on the algebraic relation between the orthogonal moments and geometric ones. Obviously, the invariance of this approach comes from the geometric moment invariants. Here, the conversion relationship is universal, since polynomials are formed directly from linear combinations of sequences $\{x^0, x^1, \ldots, x^n\}$, and harmonic functions can be written in a similar form via the *Taylor series*. Note that image normalization using geometric moments followed by an orthogonal moments–based description and the orthogonal moment invariants derived by a linear combination of geometric moment invariants play a very similar role in theory; in practice, different methods have different issues on numerical stability and computational complexity [8].
- Explicit method [88–90]: This approach seeks to derive invariants directly from the orthogonal moments. Mathematically, it tries to satisfy the identity $\mathcal{I}(\{\langle f, V_{nm}\rangle\}) = \mathcal{I}(\{\langle \mathcal{D}(f), V_{nm}\rangle\})$ by designing V_{nm} and $\mathcal{I}$. The invariance of this approach directly comes from the given moments. A very common method in this approach is to explicitly construct the rotation invariants of the circular moments. Due to the angular basis function $A_m(\theta) = \exp(\boldsymbol{j}m\theta)$ and *Fourier Shift Theorem*, the circular moments defined in Equation (1) and

Equation (4) of the rotated image $f^{\text{rot}}(r,\theta) = f(r,\theta+\phi)$ are $M'_{nm} = \exp(\boldsymbol{j}m\phi)M_{nm}$. Hence, image rotation operation only affects the phase of circular moments. The traditional method of achieving phase cancellation is based on magnitude, i.e., $|M'_{nm}| = |\exp(\boldsymbol{j}m\phi)M_{nm}| = |M_{nm}|$. However, such simple methods discard too much image information: leaving only the magnitude, and the more important phase is ignored [91, 92]. For this, Flusser [93] proposed the complex-valued rotation invariants containing phase information. Let $L \geq 1$ and $n_i \in \mathbb{N}$, $m_i, k_i \in \mathbb{Z}$, $i = 1,\ldots,L$, such that $\sum_{i=1}^{L} m_i k_i = 0$. Then, the complex-valued rotation invariants of circular moments can be defined as $inv = \prod_{i=1}^{L} [M_{n_i m_i}]^{k_i}$ [93]. In fact, the normalization method may be impractical for many applications, e.g., the dense image block representation. Moreover, the moments computed by normalization scheme may differ from the true moments of the standard image, owing to errors in the evaluation of the normalization parameters. As for the indirect method, a long time is allocated to compute the polynomial coefficients, hence requiring the recursive calculation of such coefficients. In contrast, due to its direct structure, the explicit method is more applicable in a variety of scenarios and can usually achieve higher accuracy and efficiency [62].

4.3.2 Noise. Different from the affine transformation, the noise attack acts on the intensity domain, that is, the degradation operation $\mathcal{D}$ is a mapping from the original intensity function $f(x,y)$ to the new intensity function $f'(x,y)$, $\mathcal{D} : f(x,y) \rightarrow f'(x,y)$. Mathematically, based on a specific noise function $\eta(x,y)$, the common additive noise is defined as

$$f'(x,y) = f(x,y) + \eta(x,y), \tag{24}$$

and multiplicative noise is defined as

$$f'(x,y) = f(x,y) \times \eta(x,y). \tag{25}$$

For enhancing the robustness to noise, it is difficult to adopt a similar normalization as in the affine invariant representation, because the inverse problems are often *ill-conditioned* or *ill-posed* [94, 95]. Currently, a common strategy in moment-based image representation is to convert the image to a new space with higher ***Signal-to-Noise Ratio*** (**SNR**) via a specific transformation $\mathcal{T}$, e.g., *Radon space* [96–98]. Based on this new space, V_{nm} and $\mathcal{I}$ are designed for achieving a robust representation $\mathcal{R}$:

$$\mathcal{I}(\{\langle \mathcal{T}(f), V_{nm} \rangle\}) \simeq \mathcal{I}(\{\langle \mathcal{T}(\mathcal{D}(f)), V_{nm} \rangle\}). \tag{26}$$

An image representation based on the Radon transform has the advantage of being robust to additive noise [99]. The Radon transform of an image function $f(x,y)$ is defined as

$$\text{Rad}_f(\rho,\alpha) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x,y)\delta_{\rho,\, x\cos\alpha + y\sin\alpha}\, dxdy, \tag{27}$$

where δ_{ij} is the Kronecker delta function defined in Equation (3), and $\rho = x\cos\alpha + y\sin\alpha$ is a straight line with the angle α (w.r.t. the y-axis) and the distance ρ (w.r.t. the origin). The Radon transform of the noisy image defined in Equation (24) can be written as

$$\text{Rad}_{f'}(\rho,\alpha) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} [f(x,y) + \eta(x,y)]\delta_{\rho,\, x\cos\alpha + y\sin\alpha}\, dxdy = \text{Rad}_f(\rho,\alpha) + \text{Rad}_\eta(\rho,\alpha). \tag{28}$$

In the continuous domain, the Radon transform of noise $\text{Rad}_\eta(\rho,\alpha)$ is equal to the mean value of the noise, which is assumed to be zero. Thus, we have

$$\text{Rad}_{f'}(\rho,\alpha) = \text{Rad}_f(\rho,\alpha). \tag{29}$$

This means that zero-mean additive noise has no effect on the Radon transform of the image. As for multiplicative noise as Equation (25), a potential strategy is to use logarithmic transformation as a pre-processing to separate the image and noise into two additive parts, i.e., $\log[f(x,y) \times \eta(x,y)] = \log[f(x,y)] + \log[\eta(x,y)]$. In practice, however, the perfect identity of Equation (29) does not hold because the image and noise are sampled and quantized. For the discrete case, the SNR of the image in Radon space is still significantly higher than that of the original image, meaning better robustness. Readers are referred to the papers of Hoang and Tabbone [97] and Jafari-Khouzani and Soltanian-Zadeh [99] for detailed theoretical analysis and experiments.

4.3.3 Blurring. Similar to the noise attack, the blurring also acts on the intensity domain, that is, $\mathcal{D} : f(x,y) \rightarrow f'(x,y)$. Mathematically, based on a specific ***Point-Spread Function*** (**PSF**) $\mu(x,y)$, the observed blurred image can be described as a convolution:

$$f'(x,y) = f(x,y) \otimes \mu(x,y). \tag{30}$$

Therefore, blur invariance can also be called convolution invariance. For image blurring,"normalization" means in fact blind deconvolution, which is a strongly ill-conditioned or ill-posed problem [100, 101]. At present, the invariant representation of blurred images usually relies on specific image space transformation such as projection operation similar to Radon transform. The main idea can also be expressed by Equation (26). In this regard, the works of Flusser et al. [102–105] are remarkable.

4.3.4 Recent Advance of Robustness/Invariance Optimization. The affine moment invariants of 2D and 3D images are well-established tools [106–109]. In the books of Flusser et al. [8, 16], explicit affine invariants and normalization strategy using geometric/complex moments are derived in detail. Not only that, such invariants can easily be extended to orthogonal moments through the inherent conversion relationship. Hence, the work of Flusser et al. on affine invariants is very generic and remains one of the most landmark efforts in the field. More recently, the extension of invariants has attracted strong scientific interest [110]. (1) Extension to other *transform groups*, such as the invariants under reflection [111], projective [112], Mobius [113], and chiral [114] transformations. (2) Extension to other *data dimensions* or *formats*: such as invariants of color image [115], curve [116], surface [117], vector field [118, 119], and tensor field [120–122]. (3) Extension to other *moments*, such as the invariants of OFMM [123], JFM [124], Fourier-Mellin transform [125], radial Tchebichef moments [126], radial Legendre moments [127], and Gaussian-Hermite moments [128–131]. As a noteworthy work, Li et al. [132] deeply explored the structure of invariants. They proposed two fundamental ***Generating Functions*** (**GFs**), which can encode geometric moment invariants and further construct shape descriptors, just like DNA encodes proteins. They also found that Hu's seven invariants can be further decomposed into a simpler set of ***Primitive Invariants*** (**PIs**), which may mean a new perspective of invariant study [133].

It is well known that the inverse problems of image noise and blurring are usually more difficult to solve than the inverse problems of affine transformations. This means that the normalization methods, i.e., denoising and deblurring, are very challenging tasks. Such approach is, generally, slow, and unstable due to the restoration artifacts [102]. For the brute-force path such as CNN, the lack of inherent invariance causes them to be very sensitive to noise and blurring operations not seen in the training. A recent work [134] confirmed this conclusion through extensive experiments. The most common strategy to alleviate this problem is data augmentation. However, it is very time- and memory-consuming, just"learning by rote". Due to the above facts, the moment-based invariant representation of noisy/blurred images is crucial to many practical applications. In this path, a main goal pursued by the researchers is to simplify the assumptions about degradation, which are the basis for the design of invariants. For example, by the additional constraints on PSF, they derived invariants to motion blur [135], axially symmetric blur in the case of two axes [136], circularly symmetric blur [137, 138], arbitrary N-fold symmetric blur [103], circularly symmetric Gaussian blur [104], and general (anisotropic) Gaussian blur [102]. Recently, a promising invariant representation to blurring was proposed by Kostková et al. [102]. The main contribution is the design of the invariants to general Gaussian blur, where the PSF is a Gaussian function with unknown parameters, i.e., the blur kernel may be arbitrarily oriented, scaled, and elongated.

4.4 Definition Extension

Starting from different application scenarios and optimization goals, mathematical extension on the definitions of classical moment basis functions is also a popular research topic. In the following, we will introduce some common paths of definition extension.

4.4.1 Quaternion. Mathematically, the gray-level image function $f(x,y)$ can be defined as a mapping from 2D image plane to 1D intensity value, i.e., $f:\mathbb{R}^2\to\mathbb{R}$; the color image function $\boldsymbol{f}(x,y)=\{f_R(x,y),f_G(x,y),f_B(x,y)\}$ can be defined as a mapping from 2D image plane to 3D intensity value, i.e., $\boldsymbol{f}:\mathbb{R}^2\to\mathbb{R}^3$. Here, $f_R(x,y)$, $f_G(x,y)$, and $f_B(x,y)$ are the Red, Green, and Blue components of color image in the RGB model, respectively. Note that other three-channel color models are also applicable to the analysis in this section.

In Sections 2 and 3, the listed classical moments were directly designed for gray-level images but not for color images. When dealing with color image, there are two straightforward strategies [139, 140]: (1) graying method, calculating the moments of the gray-level version of color image; and (2) channel-wise method, directly calculating the moments of each color channel. However, the graying method may lose some significant color information. In addition, the channel-wise method can hardly produce the most compact representation of a color image and ignores the correlation between different color channels.

For solving the above problems, a common strategy is to define the color image $\boldsymbol{f}(x,y)$ as a mapping from 2D image plane to *quaternion* intensity value $f_R(x,y)\boldsymbol{i}+f_G(x,y)\boldsymbol{j}+f_B(x,y)\boldsymbol{k}$, i.e., $\boldsymbol{f}:\mathbb{R}^2\to\mathbb{H}$. Correspondingly, the basis function V_{nm} is also extended from the real domain $\mathbb{R}$ or the complex domain $\mathbb{C}$ to the quaternion domain $\mathbb{H}$, i.e., new basis functions $\boldsymbol{V}_{nm}\in\mathbb{H}$, for achieving a well counterpart $\langle\boldsymbol{f},\boldsymbol{V}_{nm}\rangle$ of the original inner product $\langle f,V_{nm}\rangle$ [67, 141, 142]. For more details on quaternion algebra and quaternion moments, we encourage readers to see the papers by Chen et al. [141, 142].

4.4.2 High-Dimensional. Quaternion extension acts on the range of images and basis functions, while high-dimensional extension acts on their domain. In Sections 2 and 3, the listed classical moments were directly designed for 2D images. When dealing with images defined in high-dimensional space $\mathbb{R}^d$ (mainly 3D image, $d=3$), the domain D of the basis functions V_{nm} should also be extended from $\mathbb{R}^2$ to $\mathbb{R}^d$ [16].

For Cartesian moments, such extension is quite straightforward. Based on a 1D orthogonality polynomials set, we can generate d-dimensional orthogonal basis functions by using the same polynomials set for each of the directions/variables [145].

For circular moments, extension to high-dimensional space is more difficult. In the 3D case, a common choice is replacing the angular basis function $A_m(\theta)=\exp(\boldsymbol{j}m\theta)$ with the *spherical harmonic* $Y_{ml}(\theta,\varphi)$ of degree $l\in\mathbb{N}$ and order $m\in\mathbb{Z}$ [146, 147]:

$$Y_{ml}(\theta,\varphi)=\sqrt{\frac{(2l+1)(l-m)!}{4\pi(l+m)!}}L_{ml}(\cos\theta)\exp(\boldsymbol{j}m\varphi), \tag{31}$$

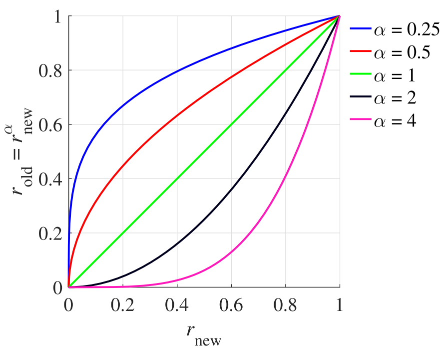


Fig. 8. Illustration of the variable substitution in fractional-order extension.

where $\theta \in [0, \pi)$ and $\varphi \in [0, 2\pi)$ are inclination and azimuth of the spherical coordinate system, respectively, $|m| \leq l$, and L_{ml} represents the associated Legendre functions, which can be written explicitly as

$$L_{ml}(x) = (-1)^m 2^l (1 - x^2)^{\frac{m}{2}} \sum_{k=m}^{l} \frac{k! x^{k-m}}{(k-m)!} \binom{l}{k} \binom{\frac{l+k+1}{2}}{l}. \tag{32}$$

The spherical harmonic $Y_{ml}(\theta, \varphi)$ satisfies the orthogonality condition:

$$\langle Y_{ml}, Y_{m'l'} \rangle = \int_0^{2\pi} \int_0^{\pi} Y_{ml}(\theta, \varphi) Y^*_{m'l'}(\theta, \varphi) \sin\theta d\theta d\varphi = \delta_{mm'} \delta_{ll'}. \tag{33}$$

Therefore, the 3D orthogonal basis functions in spherical coordinates can be constructed by combining the spherical harmonic $Y_{ml}(\theta, \varphi)$ and radial basis functions $R_n(r)$ with some slight modifications [148].

4.4.3 Fractional-Order. In Sections 2 and 3, the listed classical moments were designed to rely on integer-order domain $(n, m) \in \mathbb{Z}^2$. Recently, an interesting idea on extending the order domain of the classical moments has emerged [149–151]. It introduces a fractional-order parameter $\alpha \in \mathbb{R}$ through certain suitable variable substitutions, e.g., $r := r^\alpha \in [0, 1]$ (circular moments) or $x := x^\alpha, y := y^\alpha \in [0, 1]$ (Cartesian moments). By the substitution, the order of the newly defined moments can be extended to real domain, e.g., $\alpha n \in \mathbb{R}$ (circular moments) or $(\alpha n, \alpha m) \in \mathbb{R}^2$ (Cartesian moments).

It is worth noting that such kind of fractional-order moments is not only the mathematical extension of classical moments, but also has a distinctive *time-frequency analysis* capability [152]. Specifically, the fractional-order moments are able to control the zero distributions of the basis functions by changing the value of the fractional-order parameter. According to the research on the information suppression problem (mentioned in Section 3.4), the distribution of zeros of the basis functions is a very important property because it is closely related to the description emphasis of the moments in the spatial domain. As a result, the computed fractional-order moments are able to put emphasis on certain regions of an image, which are useful for solving information suppression issues and extracting image local features [149–151].

We will explain why fractional-order extension brings the time-frequency analysis capability. Take the fractional-order circular moments as an example, where a new variable $r_{\text{new}} \in [0, 1]$ is used in the definition with $r_{\text{old}} = {r_{\text{new}}}^{\alpha} \in [0, 1]$ and $\alpha \in \mathbb{R}^+$. As illustrated in Figure 8, we can derive the following conclusions on the distribution of zeros and the description emphasis of extracted moments [152]:

- When $\alpha = 1$, the zeros of radial basis functions and description emphasis are the same as the corresponding integer-order version due to $r_{\text{old}} = {r_{\text{new}}}^{\alpha} = r_{\text{new}}$.
- When $\alpha < 1$, the zeros of radial basis functions are biased toward 0 due to $r_{\text{old}} = {r_{\text{new}}}^{\alpha} > r_{\text{new}}$, meaning more emphasis on the inner region of the image.
- When $\alpha > 1$, the zeros of radial basis functions are biased toward 1 due to $r_{\text{old}} = {r_{\text{new}}}^{\alpha} < r_{\text{new}}$, meaning more emphasis on the outer region of the image.

4.4.4 Other. In addition to the above common paths, several other extension strategies can be found in the literature.

Zhu [153] uses *multivariate* orthogonal polynomials [154], which are the tensor product of two different orthogonal polynomials, for the definition of Cartesian moments. In contrast, the classical strategy is to use the tensor product of two same orthogonal polynomials. This new definition is more flexible and may have better performance in certain situations [155–158].

Wang et al. [159] proposed *semi-orthogonal* moments to adjust the frequency-domain nature and spatial-domain description emphasis of the basis functions. It adopts a certain modulation function (may with parameter [160, 161]) to weight the basis functions. Such semi-orthogonal moments are reported to be powerful for time-frequency analysis, but the orthogonality condition of their basis functions is not held.

Zhu et al. [162, 163] adopted some *generalized* polynomials to extend the definition of existing classical moments. The corresponding recursive calculation formulas are also explicitly given in the papers. Note that a newly introduced parameter in these methods exhibits a similar property to the fractional-order parameter (mentioned in Section 4.4.3), i.e., the ability to adjust the distribution of zeros. One possible implication is that there may be a mathematical connection between the two. Some similar papers on definition extension are [164–167].

4.4.5 Recent Advance of Definition Extension. Quaternion extension has become a very common strategy in moment-based color image processing [139, 168–172]. In addition, quaternion theory is increasingly used in related fields, including both handcrafted [173–175] and learning-based representation [176–179]. We list below some recent advances in quaternion moments. Chen et al. [180] adopted quaternion algebra for representing **RGB-D** (**RGB and depth**) images, where the real part of the quaternion number encodes the depth component. Yamni et al. [181] proposed a new category of moments for color stereo image representation, called octonion moments. Such kind of moments is a natural generalization of quaternion moments, relying on octonion algebra. We believe that further research in this field will be inspired by the advance of hyper-complex algebra such as [182].

High-dimensional extension of image moments appeared very early [148] and attracted significant attention in the last decade [145, 158, 185–189]. A possible explanation for its popularity might be the rapid development of devices and technologies related to 3D images such as medical imaging [183] and computer graphics [184]. Recent work in this field mainly focuses on the new definitions [145, 158], invariants [188, 189], and accurate/fast calculations [185–187] of 3D image moments. Note that most of the main ideas for accurate/fast calculations of 2D image moments (mentioned in Sections 4.1 and 4.2) can be naturally generalized to the 3D case. As for the 3D moment invariants, some promising works have been analyzed in Section 4.3.

Most recently, more attention of researchers has been drawn to the fractional-order extension, compared to the above two paths [190–192]. In this regard, the earliest work we find was presented by Hoang and Tabbone [149, 150]. They extend the existing harmonic function–based orthogonal moments (mentioned in Section 3.2) to a generic version, called **Generic Polar Harmonic Transforms** (**GPHT**), by fractional-order extension. A key contribution is that the time-frequency nature of GPHT and its potential in image representation were first discovered and analyzed. Xiao et al. [151] defined a general framework for the fractional-order extension of Jacobi polynomial-based orthogonal moments in both the circular and the Cartesian case (mentioned in Sections 2 and 3.1). Most of the subsequent related papers [193, 194, 194–196, 198–200] follow Hoang and Tabbone [149, 150] and Xiao et al. [151], where the main ideas of [198–200] are actually special cases of [149, 150]. Note that all the above related work focuses on solving the information suppression problem or extracting local image information. In a recent paper, Yang et al. [152] define a new set of generic fractional-order orthogonal moments, called **Fractional-order Jacobi-Fourier Moments** (**FJFM**), taking Jacobi polynomial-based classical and fractional-order orthogonal moments as special cases. More importantly, they found that all the above methods have a common defect in the image global representation. Such global representation using a specific fractional parameter (determined experimentally) only brings a slight performance improvement compared to the classic moments, mainly due to the contradiction between the robustness and discriminability. Starting from this, Yang et al. [152] take a first step to improve the performance of global representation using the time-frequency property of fractional-order orthogonal moments.

4.5 Application

In addition to the theoretical works described above, the application research of orthogonal moments is also active. Such applications cover many familiar problems of image processing, computer vision, and information security, as well as some interdisciplinary frontiers. In this section, considering the focus of this survey, we aim to draw a high-level intuition for the applications of image moments without involving technical details.

- *Image processing* – For the low-level vision tasks, image moments are popular choices. The application involves different contents of image processing: low-level feature detection and representation (e.g., detection of edge [201] and keypoint [202], representation of interest region [203], texture [55], and optical flow [204]), degraded image restoration and representation (e.g., denoising [205], deblurring [206], and superresolution [207]), image registration [208], image compression, coding and communication [209], and image quality assessment [210]. Similarly, moments have also been employed for the low-level vision tasks in video analysis [211] and computer graphics [212].
- *Visual understanding* – For the high-level vision tasks, image moments have been explored in different applications of computer vision. The related works are found in image classification [213], instance retrieval [214], object detection [215], and semantic segmentation [216]. Another major class of applications is pattern recognition, including behavior recognition [217], text recognition [218], biometric recognition [219], and sentiment recognition [220].
- *Information security* – Moments and moment invariants have shown significant impact on the research of information security, especially for the visual media. The popularity should be attributed to the robustness of moment-based representation, which is consistent with the *two-player* nature of security research. Successful applications mainly comprise digital watermarking [221], steganography [222], perceptual hashing [223], and passive media forensics (e.g., copy-move [224] and splicing [180] detection).
- *Interdisciplinary* – Since visual information processing and understanding are common tasks in many disciplines, the moment-based representation can naturally be extended to these fields. Typical interdisciplinary applications cover medicine (e.g., medical imaging [225]), geography (e.g., remote sensing [226]), robotics (e.g., visual servoing [227]), physics (e.g., optics [228] and fluid mechanics [118]), chemistry (e.g., analytical chemistry [229]), biology (e.g., protein structure representation [230]), and materials science (e.g., atomic environment representation [231]).

From the above review, two observations should be mentioned: (1) many works in Sections 4.1, 4.2, 4.3, and 4.4 are generic in their nature and are found in the above application areas; (2) moment-based representation plays a key role in a variety of scenarios that require high efficiency or strong robustness.

5 SOFTWARE PACKAGE AND EXPERIMENT RESULTS

In this section, we will give an open source software for a variety of widely used orthogonal moments. Some accurate/fast calculation, robustness/invariance optimization, and definition extension strategies are also included in the package. This software is thus called `MomentToolbox`, which is available at `https://github.com/ShurenQi/MomentToolbox`.

On this unified base, we will evaluate the accuracy/complexity, representation capability, and robustness/invariance of these methods through moment calculation, image reconstruction, and pattern recognition experiments, respectively.

Due to space limitations, the detailed settings, results, and analysis of these experiments are moved to the online Supplementary Material. We strongly encourage the reader to access this appendix via the `DOI` link of this article. The main observations and conclusions for the experiments are presented below.

— *Moment calculation: Accuracy and complexity* – (1) In the direct calculation scenario, harmonic function-based moments have better computational complexity and accuracy, compared with Jacobi polynomial-based moments and eigenfunction-based moments. (2) The easy-to-implement recursive strategy can effectively overcome the numerical stability issue of Jacobi polynomial-based moments; also with better efficiency, it should be promoted in real-world applications. (3) As a promising research path, the harmonic function-based moments calculated by FFT show superior performance in terms of both complexity and accuracy.

— *Image reconstruction: Representation capability* – (1) In the family of classical Jacobi polynomial-based methods, except for ZM and PZM, the reconstruction/representation qualities of all other methods are similar, due to their similar mathematical properties. (2) In fact, ZM and PZM require higher-order moments to obtain a close performance to other methods, but possibly at the cost of reduced robustness to signal corruptions. (3) In contrast, the harmonic function-based moments, especially EFM, PCET, and GPCET, are generally better than other types of methods in the clean image representation. (4) It is worth noting that the time-frequency analysis capability of fractional-order moments may useful for certain scenarios of image representation. For this, readers can refer to the preliminary work in [152].

— *Pattern recognition: Robustness and invariance* – (1) In general, the above phenomena and conclusions from the image reconstruction experiment are continued in this pattern recognition experiment. (2) Compared with some learning methods [4, 232–234], orthogonal moment methods exhibit certain advantages for recognizing the image variants under geometric transformations and signal corruptions. The invariance and independence of the moment-based descriptor draw a distinction between other methods, leading to potential benefits on small-scale robust recognition problems.

6 CONCLUDING REMARKS AND FUTURE DIRECTIONS

Robust and discriminative image representation is a long-lasting battle in computer vision and pattern recognition. In this article, we have presented a comprehensive survey on the orthogonal moment methods in image representation.

Starting from the review on the basic theories and classical methods in the area, we abstracted several basic conditions that an efficient moment-based representation should satisfy: the invariance/robustness to affine transformations and signal corruptions, the discriminability to a large number of patterns, and reasonable computational complexity and accuracy. Based on these observations, this article aimed to analyze the motivation and successful experiences behind the recent advances in fast/accurate calculation, robustness/invariance optimization, definition extension, and application. Note that such overall theoretical analysis of the state-of-the-art research progress is mostly ignored in previous studies.

In addition to the above theoretical contributions, we also provided extensive open source implementations and experimental evaluations at the practical level. For the first time, a software package called `MomentToolbox` is available for the image representation community, covering classical methods and improvement strategies in the field of image moments. With this software, this article has evaluated the accuracy/complexity, representation capability, and robustness/invariance of the widely used methods through moment calculation, image reconstruction, and pattern recognition experiments, respectively. To the best of our knowledge, such overall performance statistics of the state-of-the-art methods have not been given until this work.

As can be seen from this survey, over a period of nearly six decades, the widespread studies on this field have resulted in a great amount of achievements. Despite its long history, moment-based representation appears to be still in development. This is not so surprising, considering the fundamental role of image representation, i.e., the performance of computer vision and pattern recognition methods is heavily dependent on the choice of data representation. With this premise, it is important to try to identify the most promising areas for future research.

— *Moment invariant theory in bag-of-visual-words model.* As the most competitive representation model in the pre-CNN era, the **Bag-of-Visual-Words** (**BoVW**) model is a handcrafted algorithm in which local features are extracted, encoded, and summarized into global image representation. One of the main difficulties faced by the BoVW model is the unsatisfactory robustness of the representation [235, 236]. Obviously, the core of the improvement is to expand the invariance of local features (including the descriptor and detector). Therefore, the moment-based local descriptor/detector with good invariance to geometric transformations and signal corruptions is promising in alleviating this problem. Specifically, we noted a number of potential efforts [76, 92, 202, 203, 237–242].

— *Moment invariant theory in deep-learning model.* As one of the most important representation methods in deep learning, CNN serves as a hierarchical model for large-scale visual tasks. The large number of neurons contained in the network allows CNN to fit any complicated data distribution, meaning a strong discriminability. For this reason, CNN has received widespread attention recently. However, its problems are also commonly reported such as high time/space complexity and difficulty in achieving satisfactory robustness [134, 235]. In this respect, the mathematical constraints of invariance and independence behind the moment-based image representation are useful for solving these problems. We believe that such exploration of introducing knowledge into data-driven algorithms is promising. Specifically, we noted a number of potential efforts [243–250].

— *Moment invariant theory in real-world application.* In fact, the application of moments and moment invariants is still an active field. This is because different application scenarios have different requirements for the accuracy, complexity, invariance/robustness, and discriminability of image representation. In other words, the special optimization of moment-based image representation considering background knowledge is quite necessary in practice. Specifically, we noted a number of potential efforts [74, 76, 92, 118, 173, 205, 239, 240, 251].

REFERENCES

[1] C. E. Shannon. 1984. A. mathematical theory of communication. *Bell Syst. Tech. J.* 27, 3 (1984), 379–423.

[2] L. Zheng, Y. Yang, and Q. Tian. 2018. Sift meets CNN: A decade survey of instance retrieval. *IEEE Trans. Pattern Anal. Mach. Intell.* 40, 5 (2018), 1224–1244.

[3] J. Gu et al. 2018. *Pattern Recognit.* 77 (2018), 354–377.

[4] T. H. Chan, K. Jia, S. Gao, J. Lu, Z. Zeng, and Y. Ma. 2015. PCANet: A simple deep learning baseline for image classification?. *IEEE Trans. Image Process.* 24, 12 (2015), 5017–5032.

[5] K. He, X. Zhang, S. Ren, and J. Sun. 2015. Spatial pyramid pooling in deep convolutional networks for visual recognition. *IEEE Trans. Pattern Anal. Mach. Intell.* 37, 9 (2015), 1904–1916.

[6] V. Balntas, K. Lenc, A. Vedaldi, T. Tuytelaars, J. Matas, and K. Mikolajczyk. 2020. H-patches: A benchmark and evaluation of handcrafted and learned local descriptors. *IEEE Trans. Pattern Anal. Mach. Intell.* 42, 11 (2020), 2825–2841.

[7] K. Mikolajczyk and C. Schmid. 2005. A performance evaluation of local descriptors. *IEEE Trans. Pattern Anal. Mach. Intell.* 27, 10, (2005) 1615–1630.

[8] J. Flusser, B. Zitova, and T. Suk. 2009. *Moments and Moment Invariants in Pattern Recognition.* John Wiley & Sons,

[9] J. Yuan, Y. Wu, and M. Yang. 2007. Discovery of collocation patterns: From visual words to visual phrases. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 1–8.

[10] M. K. Hu. 1962. Visual pattern recognition by moment invariants. *IRE Trans. Inf. Theory* 8, 2 (1962), 179–187.

[11] R. Mukundan and K. Ramakrishnan. 1998. *Moment Functions in Image Analysis: Theory and Applications.* World Scientific.

[12] M. Pawlak. 2006. *Image Analysis by Moments: Reconstruction and Computational Aspects.* Oficyna Wydawnicza Politechniki Wrocławskiej.

[13] H. Shu, L. Luo, and J. L. Coatrieux. 2007. Moment-based approaches in imaging part 1: Basic features. *IEEE Eng. Med. Biol.* 26, 5 (2007), 70–74.

[14] H. Shu, L. Luo, and J. L. Coatrieux. 2008. Moment-based approaches in imaging part 2: Invariance. *IEEE Eng. Med. Biol.* 27, 1 (2008), 81–83.

[15] H. Shu, L. Luo, and J. L. Coatrieux. 2008. Moment-based approaches in imaging part 3: Computational considerations. *IEEE Eng. Med. Biol.* 27, 3 (2008), 89–91.

[16] J. Flusser, T. Suk, and B. Zitová. 2016. *2D and 3D Image Analysis by Moments.* John Wiley & Sons,

[17] T. V. Hoang, *Image Representations for Pattern Recognition.* Ph.D. Dissertation. Department of Computational Science. Nancy, France.

[18] G. A. Papakostas. 2014. Over 50 years of image moments and moment invariants. *Moments and Moment Invariants - Theory and Applications*, G. A. Papakostas (Ed.), Science Gate, 3–32.

[19] P. Kaur, H. S. Pannu, and A. K. Malhi. 2019. Comprehensive study of continuous orthogonal moments—A systematic review. *ACM Comput. Surv.* 52, 4 (2019).

[20] C. H. Teh and R. T. Chin. 1988. On image analysis by the methods of moments. *IEEE Trans. Pattern Anal. Mach. Intell.* 10, 4 (1988), 496–513.

[21] Y. S. Abu-Mostafa and D. Psaltis. 1984. Recognitive aspects of moment invariants. *IEEE Trans. Pattern Anal. Mach. Intell.* 6, 6 (1984), 698–706.

[22] D. Zhang and G. Lu. 2002. Shape-based image retrieval using generic Fourier descriptor, signal process. *Image Commun.* 17, 10 (2002), 825–848.

[23] R. Mukundan, S. H. Ong, and P. A. Lee. 2001. Image analysis by Tchebichef moments. *IEEE Trans. Image Process.* 10, 9 (2001), 1357–1364.

[24] P. T. Yap, R. Paramesran, and S. H. Ong. 2003. Image analysis by Krawtchouk moments. *IEEE Trans. Image Process.* 12, 11 (2003), 1367–1377.

[25] P. T. Yap, R. Paramesran, and S. H. Ong. 2007. Image analysis using Hahn moments. *IEEE Trans. Pattern Anal. Mach. Intell.* 29, 11 (2007), 2057–2062.

[26] H. Zhu, H. Shu, J. Zhou, L. Luo, and J. L. Coatrieux. 2007. Image analysis by discrete orthogonal dual Hahn moments. *Pattern Recognit. Lett.* 28, 13 (2007), 1688–1704.

[27] H. Zhu, H. Shu, J. Liang, L. Luo, and J. L. Coatrieux. 2007. Image analysis by discrete orthogonal Racah moments. *Signal Process.* 87, 4 (2007), 687–708.

[28] A. B. Bhatia and E. Wolf. 1954. On the circle polynomials of Zernike and related orthogonal sets. *Math. Proc. Camb. Philos. Soc.,* (1954), 40–48.

[29] M. R. Teague. 1980. Image analysis via the general theory of moments. *J. Opt. Soc. Am. A* 70, 8 (1980), 920–950.

[30] J. Shen. 1997. Orthogonal Gaussian-Hermite moments for image characterization. *Proc. SPIE Intell. Rob. Comput. Vis.* (1997), 224–233.

[31] K. M. Hosny. 2011. Image representation using accurate orthogonal Gegenbauer moments. *Pattern Recognit. Lett.* 32, 6 (2011), 795–804.

[32] H. Zhu. 2012. Image representation using separable two-dimensional continuous and discrete orthogonal moments. *Pattern Recognit.* 45, 4 (2012), 1540–1558.

[33] Y. Sheng and L. Shen. 1994. Orthogonal Fourier-Mellin moments for invariant pattern recognition. *J. Opt. Soc. Am. A* 11, 6 (1994), 1748–1757.

[34] Z. Ping, R. Wu, and Y. Sheng. 2002. Image description with Chebyshev-Fourier moments. *J. Opt. Soc. Am. A* 19, 9 (2002), 1748–1754.

[35] G. Amu, S. Hasi, X. Yang, and Z. Ping. 2004. Image analysis by pseudo-Jacobi (p = 4, q = 3)-Fourier moments. *Appl. Opt.* 43, 10 (2004), 2093–2101.

[36] Z. Ping, H. Ren, J. Zou, Y. Sheng, and W. Bo. 2007. Generic orthogonal moments: Jacobi-Fourier moments for invariant image description. *Pattern Recognit.* 40, 4 (2007), 1245–1254.

[37] T. V. Hoang and S. Tabbone. 2013. Errata and comments on generic orthogonal moments: Jacobi-Fourier moments for invariant image description. *Pattern Recognit.* 46, 11 (2013), 3148–3155.

[38] H. Ren, Z. Ping, W. Bo, W. Wu, and Y. Sheng. 2003. Multidistortion-invariant image recognition with radial harmonic Fourier moments. *J. Opt. Soc. Am. A* 20, 4 (2003), 631–637.

[39] H. Hu, Y. Zhang, C. Shao, and Q. Ju. 2014. Orthogonal moments based on exponent functions: Exponent-Fourier moments. *Pattern Recognit.* 47, 8 (2014), 2596–2606.

[40] P. T. Yap, X. Jiang, and A. C. Kot. 2010. Two-dimensional polar harmonic transforms for invariant image representation. *IEEE Trans. Pattern Anal. Mach. Intell.* 32, 7 (2010), 1259–1270.

[41] B. Xiao, J. Ma, and X. Wang. 2010. Image analysis by Bessel-Fourier moments. *Pattern Recognit.* 43, 8 (2010), 2620–2629.

[42] S. Liao and M. Pawlak. 1998. On the accuracy of Zernike moments for image analysis. *IEEE Trans. Pattern Anal. Mach. Intell.* 20, 12 (1998), 1358–1364.

[43] C. W. Chong, P. Raveendran, and R. Mukundan. 2003. A comparative analysis of algorithms for fast computation of Zernike moments. *Pattern Recognit.* 36, 3 (2003), 731–742.

[44] C. Y. Wee and R. Paramesran. 2007. On the computational aspects of Zernike moments. *Image Vis. Comput.* 25, 6 (2007), 967–980.

[45] C. Singh and E. Walia. 2011. Algorithms for fast computation of Zernike moments and their numerical stability. *Image Vis. Comput.* 29, 4 (2011), 251–259.

[46] X. Wang and S. Liao. 2013. Image reconstruction from orthogonal Fourier-Mellin moments. In *Proceedings of the International Conference on Image Analysis and Recognition*, 687–694.

[47] M. Nwali and S. Liao. 2019. A new fast algorithm to compute continuous moments defined in a rectangular region. *Pattern Recognit.* 89 (2019), 151–160.

[48] A. Prata and W. V. Rusch. 1989. Algorithm for computation of Zernike polynomials expansion coefficients. *Appl. Opt.* 28, 4 (1989), 749–54.

[49] C. Singh and R. Upneja. 2014. Accurate calculation of high order pseudo-Zernike moments and their numerical stability. *Digit. Signal Prog.* 27 (2014), 95–106.

[50] C. Singh and R. Upneja. 2012. Accurate computation of orthogonal Fourier-Mellin moments. *J. Math. Imaging Vis.* 44, 3 (2012), 411–431.

[51] Y. Xin, M. Pawlak, and S. Liao. 2007. Accurate computation of Zernike moments in polar coordinates. *IEEE Trans. Image Process.* 16, 2 (2007), 581–587.

[52] K. M. Hosny, M. A. Shouman, and H. M. A. Salam. 2011. Fast computation of orthogonal Fourier-Mellin moments in polar coordinates. *J. Real-Time Image Process.* 6, 2 (2011), 73–80.

[53] C. Camacho-Bello, C. Toxqui-Quitl, A. Padilla-Vivanco, and J. J. Baez-Rojas. 2014. High-precision and fast computation of Jacobi-Fourier moments for image description. *J. Opt. Soc. Am. A* 31, 1 (2014), 124–134.

[54] R. Mukundan and K. R. Ramakrishnan. 1995. Fast computation of Legendre and Zernike moments. *Pattern Recognit.* 28, 9 (1995), 1433–1475.

[55] C. D. Ruberto, L. Putzu, and G. Rodriguez. 2018. Fast and accurate computation of orthogonal moments for texture analysis. *Pattern Recognit.* 83 (2018), 498–510.

[56] R. Upneja and C. Singh. 2015. Fast computation of Jacobi-Fourier moments for invariant image recognition. *Pattern Recognit.* 48, 5 (2015), 1836–1843.

[57] G. A. Papakostas, Y. S. Boutalis, D. A. Karras, and B. G. Mertzios. 2007. A new class of Zernike moments for computer vision applications. *Inf. Sci.* 177, 13 (2007), 2802–2819.

[58] G. A. Papakostas, Y. S. Boutalis, D. A. Karras, and B. G. Mertzios. 2009. Modified factorial-free direct methods for Zernike and pseudo-Zernike moment computation. *IEEE Trans. Instrum. Meas.* 58, 7 (2009), 2121–2131.

[59] J. Saez-Landete. 2017. Comments on fast computation of Jacobi-Fourier moments for invariant image recognition. *Pattern Recognit.* 67 (2017), 16–22.
[60] K. M. Hosny and M. M. Darwish. 2019. A kernel-based method for fast and accurate computation of PHT in polar coordinates. *J. Real-Time Image Process.* 16, 4 (2019), 1235–1247.
[61] A. Averbuch, R. R. Coifman, D. L. Donoho, M. Elad, and M. Israeli. 2006. Fast and accurate polar Fourier transform. *Appl. Comput. Harmon. Anal.* 21, 2 (2006), 145–167.
[62] H. Yang, S. Qi, C. Wang, S. Yang, and X. Wang. 2020. Image analysis by log-polar exponent-Fourier moments. *Pattern Recognit.* 101 (2020).
[63] T. V. Hoang and S. Tabbone. 2014. Fast generic polar harmonic transforms. *IEEE Trans. Image Process.* 23, 7 (2014), 2961–2971.
[64] C. Singh and S. K. Ranade. 2013. A. high capacity image adaptive watermarking scheme with radial harmonic Fourier moments. *Digit. Signal Prog.* 23, 5 (2013), 1470–1482.
[65] T. V. Hoang and S. Tabbone. 2012. Fast computation of orthogonal polar harmonic transforms. In *Proceedings of the International Conference on Pattern Recognition*, 3160–3163.
[66] C. Singh and A. Kaur. 2015. Fast computation of polar harmonic transforms. *J. Real-Time Image Process/* 10, 1 (2015), 59–66.
[67] H. Yang, S. Qi, P. Niu, and X. Wang. 2020. Color image zero-watermarking based on fast quaternion generic polar complex exponential transform, signal process. *Image Commun.* 82 (2020).
[68] C. Wang, X. Wang, and Z. Xia. 2016. Geometrically invariant image watermarking based on fast radial harmonic Fourier moments. *Signal Process. Image Commun.* 45 (2016), 10–23.
[69] X. Wang, C. Wang, H. Yang, and P. Niu. 2021. Robust and effective multiple copy-move forgeries detection and localization. *Pattern Anal. Appl.* 24 (2021), 1025–1046.
[70] Z. Ping, Y. Jiang, S. Zhou, and Y. Wu. 2014. FFT algorithm of complex exponent moments and its application in image recognition. In *Proceedings of the SPIE International Conference on Digital Image Processing*, 4177–4180.
[71] S. P. Singh and S. Urooj. 2017. Accurate and fast computation of exponent-Fourier moment. *Arab. J. Sci. Eng.* 42, 8 (2017), 3299–3306.
[72] S. K. Hwang and W. Y. Kim. 2006. A. novel approach to the fast computation of Zernike moments. *Pattern Recognit.* 39, 11 (2006), 2065–2076.
[73] M. S. Al-Rawi. 2010. Fast computation of pseudo Zernike moments. *J. Real-Time Image Process.* 5, 1 (2010), 3–10.
[74] B. Chen, G. Coatrieux, J. Wu, Z. Dong, J. L. Coatrieux, and H. Shu. 2015. Fast computation of sliding discrete Tchebichef moments and its application in duplicated regions detection. *IEEE Trans. Signal Process.* 63, 20 (2015), 5424–5436.
[75] J. Martinez and F. Thomas. 2002. Efficient computation of local geometric moments. *IEEE Trans. Image Process.* 11, 9 (2002), 1102–1111.
[76] A. Bera, P. Klesk, and D. Sychel. 2019. Constant-time calculation of Zernike moments for detection with rotational invariance. *IEEE Trans. Pattern Anal. Mach. Intell.* 41, 3 (2019), 537–551.
[77] H. Bay, T. Tuytelaars, and L. V. Gool. 2006. SURF: Speeded up robust features. In *Proceedings of the European Conference on Computer Vision*, 404–417.
[78] R. Benouini, I. Batioua, K. Zenkouar, A. Zahi, H. E. Fadili, and H. Qjidaa. 2019. Fast and accurate computation of Racah moment invariants for image classification. *Pattern Recognit.* 91 (2019), 100–110.
[79] S. Pei and C. Lin. 1995. Image normalization for pattern recognition. *Image Vis. Comput.* 13, 10 (1995), 711–723.
[80] S. Dinggang and H. H. S. Ip. 1997. Generalized affine invariant image normalization. *IEEE Trans. Pattern Anal. Mach. Intell.* 19, 5 (1997), 431–471.
[81] S. Dongseok, J. K. Pollard, and J. P. Muller. 1997. Accurate geometric correction of ATSR images. *IEEE Trans. Geosci. Remote Sensing* 35, 4 (1997), 997–1006.
[82] C. Wang, X. Wang, C. Zhang, and Z. Xia. 2017. Geometric correction based color image watermarking using fuzzy least squares support vector machine and Bessel K form distribution. *Signal Process.* 134 (2017), 197–208.
[83] L. G. Brown. 1992. A. survey of image registration techniques. *ACM Comput. Surv.* 24, 4 (1992), 325–376.
[84] B. Zitova and J. Flusser. 2003. Image registration methods: A survey. *Image Vis. Comput.* 21, 11 (2003), 977–1000.
[85] C. W. Chong, P. Raveendran, and R. Mukundan. 2003. Translation invariants of Zernike moments. *Pattern Recognit.* 36, 8 (2003), 1765–1773.
[86] E. G. Karakasis, G. A. Papakostas, D. E. Koulouriotis, and V. D. Tourassis. 2013. Generalized dual Hahn moment invariants. *Pattern Recognit.* 46, 7 (2013), 1998–2014.
[87] J. Flusser and T. Suk. 1993. Pattern recognition by affine moment invariants. *Pattern Recognit.* 26, 1 (1993), 167–174.
[88] C. W. Chong, P. Raveendran, and R. Mukundan. 2004. Translation and scale invariants of Legendre moments. *Pattern Recognit.* 37, 1 (2004), 119–129.

[89] S. Belkasim, E. Hassan, and T. Obeidi. 2007. Explicit invariance of Cartesian Zernike moments. *Pattern Recognit. Lett.* 28, 15 (2007), 1969–1980.

[90] H. Zhu, H. Shu, T. Xia, L. Luo, and J. L. Coatrieux. 2007. Translation and scale invariants of Tchebichef moments. *Pattern Recognit.* 40, 9 (2007), 2530–2542.

[91] A. V. Oppenheim and J. S. Lim. 1981. The importance of phase in signals. *Proc. IEEE* 69, 5 (1981), 529–541.

[92] J. Revaud, G. Lavoue, and A. Baskurt. 2009. Improving Zernike moments comparison for optimal similarity and rotation angle retrieval. *IEEE Trans. Pattern Anal. Mach. Intell.* 31, 4 (2009), 627–636.

[93] J. Flusser. 2000. On the independence of rotation moment invariants. *Pattern Recognit.* 33, 9 (2000), 1405–1410.

[94] L. Shao, R. Yan, X. Li, and Y. Liu. 2014. From heuristic optimization to dictionary learning: A review and comprehensive comparison of image denoising algorithms. *IEEE T. Cybern.* 44, 7 (2014), 1001–1013.

[95] K. Zhang, W. Zuo, Y. Chen, D. Meng, and L. Zhang. 2017. Beyond a Gaussian denoiser: Residual learning of deep CNN for image denoising. *IEEE Trans. Image Process.* 26, 7 (2017), 3142–3155.

[96] B. Xiao, J. Cui, H. Qin, W. Li, and G. Wang. 2015. Moments and moment invariants in the radon space. *Pattern Recognit.* 48, 9 (2015), 2772–2784.

[97] T. V. Hoang and S. Tabbone. 2012. Invariant pattern recognition using the RFM descriptor. *Pattern Recognit.* 45, 1 (2012), 271–284.

[98] Q. Miao, J. Liu, W. Li, J. Shi, and Y. Wang. 2012. Three novel invariant moments based on radon and polar harmonic transforms. *Opt. Commun.* 285, 6 (2012), 1044–1048.

[99] K. Jafari-Khouzani and H. Soltanian-Zadeh. 2005. Rotation-invariant multiresolution texture analysis using radon and wavelet transforms. *IEEE Trans. Image Process.* 14, 6 (2005), 783–795.

[100] D. Kundur and D. Hatzinakos. 1996. Blind image deconvolution. *IEEE Signal Process. Mag.* 13, 3 (1996), 43–64.

[101] P. Campisi and K. Egiazarian. 2017. *Blind Image Deconvolution: Theory and Applications.* CRC Press,

[102] J. Kostkova, J. Flusser, M. Lebl, and M. Pedone. 2020. Handling Gaussian blur without deconvolution. *Pattern Recognit.* 103 (2020).

[103] J. Flusser, T. Suk, J. Boldys, and B. Zitova. 2015. Projection operators and moment invariants to image blurring. *IEEE Trans. Pattern Anal. Mach. Intell.* 37, 4 (2015), 786–802.

[104] J. Flusser, S. Farokhi, C. Hoschl, T. Suk, B. Zitova, and M. Pedone. 2016. Recognition of images degraded by Gaussian blur. *IEEE Trans. Image Process.* 25, 2 (2016), 790–806.

[105] M. Pedone, J. Flusser, and J. Heikkila. 2015. Registration of images with N-fold dihedral blur. *IEEE Trans. Image Process.* 24, 3 (2015), 1036–1045.

[106] T. Suk and J. Flusser. 2011. Tensor method for constructing 3D moment invariants. In *Proceedings of the International Conference on Computer Analysis of Images and Patterns*, 212–219.

[107] T. Suk and J. Flusser. 2004. Graph method for generating affine moment invariants. In *Proceedings of the International Conference on Pattern Recognition*, 192–195.

[108] C. H. Lo and H. S. Don. 1989. 3-D moment forms: their construction and application to object identification and positioning. *IEEE Trans. Pattern Anal. Mach. Intell.* 11, 10 (1989), 1053–1064.

[109] G. Mamistvalov. 1998. N-dimensional moment invariants and conceptual mathematical theory of recognition N-dimensional solids. *IEEE Trans. Pattern Anal. Mach. Intell.* 20, 8 (1998), 819–831.

[110] D. Xu and H. Li. 2008. Geometric moment invariants. *Pattern Recognit.* 41, 1 (2008), 240–249.

[111] E. Li and H. Li. 2017. *Reflection Invariant and Symmetry Detection.* Retrieved May 2017 from https://arxiv.org/abs/1705.10768

[112] E Li, H. Mo, D. Xu, and H. Li. 2019. Image projective invariants. *IEEE Trans. Pattern Anal. Mach. Intell.* 41, 5 (2019), 1144–1157.

[113] Z. He, M. Hanlin, H. You, L. Qi, and L. Hua. 2018. Differential and integral invariants under Mobius transformation. *Proc. Chinese Conf. Pattern Recognit. Comput. Vis.* (2018), 280–291.

[114] H. Zhang, H. Mo, Y. Hao, Q. Li, S. Li, and H. Li. 2019. Fast and efficient calculations of structural invariants of chirality. *Pattern Recognit. Lett.* 128 (2019), 270–277.

[115] H. You, M. Hanlin, L. Qi, Z. He, and L. Hua. 2019. *Dual Affine Moment Invariants.* Retrieved November 2019 from https://arxiv.org/abs/1911.08233.

[116] D. Xu and H. Li. 2006. 3-D curve moment invariants for curve recognition. In *Proceedings of Intelligent Computing in Signal Processing and Pattern Recognition*, 572–577.

[117] D. Xu and H. Li. 2006. 3-D surface moment invariants. In *Proceedings of the International Conference on Pattern Recognit*ion, 173–176.

[118] J. Kostkova, T. Suk, and J. Flusser. 2019. Affine invariants of vector fields. *IEEE Trans. Pattern Anal. Mach. Intell.* (2019).

[119] J. Kostkova, T. Suk, and J. Flusser. 2018. Affine moment invariants of vector fields. In *Proceedings of the International Conference on Image Processing*, 1338–1342.

[120] M. Langbein and H. Hagen. 2009. A generalization of moment invariants on 2D vector fields to tensor fields of arbitrary order and dimension. In *Proceedings of the International Symposium on Visual Computing*, 1151–1160.

[121] R. Bujack, J. Kasten, I. Hotz, G. Scheuermann, and E. Hitzer. 2015. Moment invariants for 3D flow fields via normalization. In *Proceedings of the IEEE Pacific Visualization Symposium*, 9–16.

[122] R. Bujack and H. Hagen. 2017. Moment invariants for multi-dimensional data. In *Modelling, Analysis, and Visualization of Anisotropy*, E. Ozerslan, T. Schultz, and I. Hotz, Eds. Mathematica and Visualization, pp. 43–64.

[123] H. Zhang, H. Shu, P. Haigron, B. Li, and L. Luo. 2010. Construction of a complete set of orthogonal Fourier-Mellin moment invariants for pattern recognition applications. *Image Vis. Comput.* 28, 1 (2010), 38–44.

[124] B. Xiao and G. Wang. 2013. Generic radial orthogonal moment invariants for invariant image recognition. *J. Vis. Commun. Image Represent.* 24, 7 (2013), 1002–1008.

[125] J. Yang, Z. Lu, Y. Y. Tang, Z. Yuan, and Y. Chen. 2020. Quasi Fourier-Mellin transform for affine invariant features. *IEEE Trans. Image Process.* 29 (2020), 4114–4129.

[126] B. Xiao, J. Ma, and J. Cui. 2012. Radial Tchebichef moment invariants for image recognition. *J. Vis. Commun. Image Represent.* 23, 2 (2012), 381–386.

[127] B. Xiao, G. Wang, and W. Li. 2014. Radial shifted Legendre moments for image analysis and invariant image recognition. *Image Vis. Comput.* 32, 12 (2014), 994–1006.

[128] B. Yang, J. Flusser, and T. Suk. 2015. Design of high-order rotation invariants from Gaussian-Hermite moments. *Signal Process.* 113 (2015), 61–67.

[129] B. Yang, J. Kostkova, J. Flusser, and T. Suk. 2017. Scale invariants from Gaussian-Hermite moments. *Signal Process.* 132 (2017), 77–84.

[130] B. Yang, J. Flusser, and J. Kautsky. 2018. Rotation of 2D orthogonal polynomials. *Pattern Recognit. Lett.* 102 (2018), 44–49.

[131] B. Yang, T. Suk, J. Flusser, Z. Shi, and X. Chen. 2018. Rotation invariants from Gaussian-Hermite moments of color images. *Signal Process.* 143 (2018), 282–291.

[132] E Li, Y. Huang, D. Xu, and H. Li. 2017. *Shape DNA: Basic Generating Functions for Geometric Moment Invariants.* Retrieved March 2017 from https://arxiv.org/abs/1703.02242

[133] E. Li and H. Li. 2017. *Isomorphism between Differential and Moment Invariants Under Affine Transform.* Retrieved May 2017 from https://arxiv.org/abs/1705.08264

[134] Y. Pei, Y. Huang, Q. Zou, X. Zhang, and S. Wang. 2019. Effects of image degradation and degradation removal to CNN-based image classification. *IEEE Trans. Pattern Anal. Mach. Intell.* (2019).

[135] J. Flusser, T. Suk, and S. Saic. 1996. Recognition of images degraded by linear motion blur without restoration. In *Theoretical Foundations of Computer Vision*, (W. Kropatsch, R. Klette, F. Solina, and R. Albrecht, Eds. Computing Supplement, 37–51.

[136] J. Flusser, T. Suk, and S. Saic. 1995. Image features invariant with respect to blur. *Pattern Recognit.* 28, 11 (1995), 1723–1755.

[137] J. Flusser and B. Zitova. 2004. Invariants to convolution with circularly symmetric PSF. 2 (2004), 11–14.

[138] B. Chen, H. Shu, H. Zhang, G. Coatrieux, L. Luo, and J. L. Coatrieux. 2011. Combined invariants to similarity transformation and to blur using orthogonal Zernike moments. *IEEE Trans. Image Process.* 20, 2 (2011), 345–360.

[139] E. G. Karakasis, G. A. Papakostas, D. E. Koulouriotis, V. D. Tourassis, and A. 2014. Unified methodology for computing accurate quaternion color moments and moment invariants. *IEEE Trans. Image Process.* 23, 2 (2014), 596–611.

[140] K. M. Hosny and M. M. Darwish. 2019. New set of multi-channel orthogonal moments for color image representation and recognition. *Pattern Recognit.* 88 (2019), 153–173.

[141] B. Chen, H. Shu, H. Zhang, G. Chen, C. Toumoulin, J. L. Dillenseger, and L. Luo. 2012. Quaternion Zernike moments and their invariants for color image analysis and object recognition. *Signal Process.* 92, 2 (2012), 308–318.

[142] B. Chen, H. Shu, G. Coatrieux, G. Chen, X. Sun, and J. L. Coatrieux. 2015. Color image analysis by quaternion-type moments. *J. Math. Imaging Vis.* 51, 1 (2015), 124–144.

[143] Y. Li. 2013. Quaternion polar harmonic transforms for color images. *IEEE Signal Process. Lett.* 20, 8 (2013), 803–806.

[144] L. Guo and M. Zhu. 2011. Quaternion Fourier-Mellin moments for color images. *Pattern Recognit.* 44, 2 (2011), 187–195.

[145] I. Batioua, R. Benouini, K. Zenkouar, S. Najah, H. E. Fadili, and H. Qjidaa. 2019. 3D image representation using separable discrete orthogonal moments. *Procedia Comput. Sci.* 148 (2019), 389–398.

[146] M. Kazhdan, T. Funkhouser and S. Rusinkiewicz. 2003. Rotation invariant spherical harmonic representation of 3D shape descriptors. *Proc. Symp. Geom. Process.* (2003), 156–164.

[147] M. Novotni and R. Klein. 2004. Shape retrieval using 3D Zernike descriptors. *Comput.-Aided Des.* 36, 11 (2004), 1047–1062.

[148] N. Canterakis. 1999. 3D Zernike moments and Zernike affine invariants for 3D image analysis and recognition. In *Proceedings of the Scandinavian Conference on Image Analysis*, 85–93.

[149] T. V. Hoang and S. Tabbone. 2011. Generic polar harmonic transforms for invariant image description. In *Proceedings of the International Conference on Image Processing*, 829–832.

[150] T. V. Hoang and S. Tabbone. 2014. Generic polar harmonic transforms for invariant image representation. *Image Vis. Comput.* 32, 8 (2014), 497–509.

[151] B. Xiao, L. Li, Y. Li, W. Li, and G. Wang. 2017. Image analysis by fractional-order orthogonal moments. *Inf. Sci.* 382 (2017), 135–149.

[152] H. Yang, S. Qi, J. Tian, P. Niu, and X. Wang. 2021. Robust and discriminative image representation: Fractional-order Jacobi-Fourier moments. *Pattern Recognit.* 115 (2021).

[153] H. Zhu. 2012. Image representation using separable two-dimensional continuous and discrete orthogonal moments. *Pattern Recognit.* 45, 4 (2012), 1540–1558.

[154] C. F. Dunkl and Y. Xu. 2014. *Orthogonal Polynomials of Several Variables.* Cambridge University Press,

[155] I. Batioua, R. Benouini, K. Zenkouar, and H. E. Fadili. 2017. Image analysis using new set of separable two-dimensional discrete orthogonal moments based on Racah polynomials.

[156] A. Hmimid, M. Sayyouri, and H. Qjidaa. 2015. Fast computation of separable two-dimensional discrete invariant moments for image classification. *Pattern Recognit.* 48, (2015), 509–521.

[157] E. D. Tsougenis, G. A. Papakostas, and D. E. Koulouriotis. 2015. Image watermarking via separable moments. *Multimed. Tools Appl.* 74, 11 (2015), 3985–4012.

[158] I. Batioua, R. Benouini, K. Zenkouar, A. Zahi, and E. F. Hakim. 2017. 3D image analysis by separable discrete orthogonal moments based on Krawtchouk and Tchebichef polynomials. *Pattern Recognit.* 71 (2017), 264–277.

[159] X. Wang, T. Yang, and F. Guo. 2016. Image analysis by circularly semi-orthogonal moments. *Pattern Recognit.* 49 (2016), 226–236.

[160] B. He, J. Cui, B. Xiao, and Y. Peng. 2019. General semi-orthogonal moments with parameter modulation. *J. Image Graph.* 24, 10 (2019), 1711–1727.

[161] B. He, J. Cui, B. Xiao, and Y. Peng. 2019. Image analysis using modified exponent-Fourier moments. *EURASIP J. Image Video Process.* 72 (2019),

[162] H. Zhu, Y. Yang, X. Zhu, Z. Gui, and H. Shu. 2014. General form for obtaining unit disc-based generalized orthogonal moments. *IEEE Trans. Image Process.* 23, 12 (2014), 5455–5469.

[163] H. Zhu, Y. Yang, Z. Gui, Y. Zhu, and Z. Chen. 2016. Image analysis by generalized Chebyshev-Fourier and generalized pseudo-Jacobi-Fourier Moments. *Pattern Recognit.* 51 (2016), 1–11.

[164] T. Xia, H. Zhu, H. Shu, P. Haigron, and L. Luo. 2007. Image description with generalized pseudo-Zernike moments. *J. Opt. Soc. Am. A* 24, 1 (2007), 50–59.

[165] M. Qi, B. Li, and H. Sun. 2016. Image representation by harmonic transforms with parameters in SL(2, R). *J. Vis. Commun. Image Represent.* 35 (2016), 184–192.

[166] B. He and J. Cui. 2019. Weighted spherical Bessel-Fourier image moments. *Cluster Comput.* 22 (2019), 12985–12996.

[167] X. Liu, Y. Wu, Z. Shao, and J. Wu. 2020. The modified generic polar harmonic transforms for image representation. *Pattern Anal. Appl.* 23, 2 (2020), 785–795.

[168] T. Yang, J. Ma, Y. Miao, X. Wang, B. Xiao, B. He, and Q. Meng. 2019. Quaternion weighted spherical Bessel-Fourier moment and its invariant for color image reconstruction and object recognition. *Inf. Sci.* 505 (2019), 388–405.

[169] C. Wang, X. Wang, Y. Li, Z. Xia, and C. Zhang. 2018. Quaternion polar polar harmonic Fourier moments for color images. *Inf. Sci.* 450 (2018), 141–156.

[170] X. Wang, W. Li, H. Yang, P. Niu, and Y. Li. 2015. Invariant quaternion radial harmonic Fourier moments for color image retrieval. *Opt. Laser Technol.* 66 (2015), 78–88.

[171] X. Wang, P. Niu, H. Yang, C. Wang, and A. Wang. 2014. A. new robust color image watermarking using local quaternion exponent moments. *Inf. Sci.* 277 (2014), 731–754.

[172] Z. Shao, H. Shu, J. Wu, B. Chen, and J. L. Coatrieux. 2014. Quaternion Bessel-Fourier moments and their invariant descriptors for object reconstruction and recognition. *Pattern Recognit.* 47, 2 (2014), 603–611.

[173] J. Wang, T. Li, X. Luo, Y. Q. Shi, and S. K. Jha. 2019. Identifying computer generated images based on quaternion central moments in color quaternion wavelet domain. *IEEE Trans. Circuits Syst. Video Technol.* 29, 9 (2019), 2775–2785.

[174] S. Said, N. L. Bihan, and S. J. Sangwine. 2008. Fast complexified quaternion Fourier transform. *IEEE Trans. Signal Process.* 56, 4 (2008), 1522–1531.

[175] T. Bulow and G. Sommer. 2001. Hypercomplex signals—A novel extension of the analytic signal to the multidimensional case. *IEEE Trans. Signal Process.* 49, 11 (2001), 2844–2852.

[176] Y. Chen, X. Xiao, and Y. Zhou. 2020. Low-rank quaternion approximation for color image processing. *IEEE Trans. Image Process.* 29 (2020), 1426–1439.

[177] S. Zhang, Y. Tay, L. Yao, and Q. Liu. 2019. Quaternion knowledge graph embeddings. In *Proceedings of Advances in Neural Information Processing Systems*, 2731–2741.

[178] Y. Liu, Y. Zheng, J. Lu, J. Cao, and L. Rutkowski. 2020. Constrained quaternion-variable convex optimization: A quaternion-valued recurrent neural network approach. *IEEE Trans. Neural Netw. Learn. Syst.* 31, 3 (2020), 1022–1035.
[179] X. Zhu, Y. Xu, H. Xu, and C. Chen. 2018. Quaternion convolutional neural networks. In *Proceedings of the European Conference on Computer Vision*, 645–661.
[180] B. Chen, X. Qi, X. Sun, and Y. Q. Shi. 2017. Quaternion pseudo-Zernike moments combining both of RGB information and depth information for color image splicing detection. *J. Vis. Commun. Image Represent.* 49 (2017), 283–290.
[181] M. Yamni, H. Karmouni, M. Sayyouri, H. Qjidaa, and J. Flusser. 2021. Novel octonion moments for color stereo image analysis. *Digit. Signal Process* 108 (2021).
[182] D. Xu and D. P. Mandic. 2015. The theory of quaternion matrix derivatives. *IEEE Trans. Signal Process.* 63, 6 (2015), 1543–1556.
[183] P. Suetens. 2017. *Fundamentals of Medical Imaging*. Cambridge University Press,
[184] S. Marschner and P. Shirley. 2018. *Fundamentals of Computer Graphics*. CRC Press.
[185] J. M. Pozo, M. C. Villa-Uriol, and A. F. Frangi. 2011. Efficient 3Dd geometric and Zernike moments computation from unstructured surface meshes. *IEEE Trans. Pattern Anal. Mach. Intell.* 33, 3 (2011), 471–484.
[186] H. Karmouni, T. Jahid, M. Sayyouri, R. E. Alami, and H. Qjidaa. 2020. Fast 3D image reconstruction by cuboids and 3D Charlier's moments. *J. Real-Time Image Process.* 17, 4 (2020), 949–965.
[187] A. Daoui, M. Yamni, O. E. Ogri, H. Karmouni, M. Sayyouri, and H. Qjidaa. 2020. New algorithm for large-sized 2D and 3D image reconstruction using higher-order Hahn moments. *Circuits Syst. Signal Process.* 39, 9 (2020), 4552–4577.
[188] R. Benouini, I. Batioua, K. Zenkouar, S. Najah, and H. Qjidaa. 2018. Efficient 3D object classification by using direct Krawtchouk moment invariants. *Multimed. Tools Appl.* 77, 20 (2018), 27517–27542.
[189] I. Batioua, R. Benouini, and K. Zenkouar. 2020. Image recognition using new set of separable three-dimensional discrete orthogonal moment invariants. *Multimed. Tools Appl.* 79 (2020), 13217–13245.
[190] B. Xiao, J. Luo, X. Bi, W. Li, and B. Chen. 2020. Fractional discrete Tchebyshev moments and their applications in image encryption and watermarking. *Inf. Sci.* 516 (2020), 545–559.
[191] M. Yamni, A. Daoui, O. E. Ogri, H. Karmouni, M. Sayyouri, H. Qjidaa, and J. Flusser. 2020. Fractional Charlier moments for image reconstruction and image watermarking. *Signal Process.* 171 (2020).
[192] B. Chen, M. Yu, Q. Su, and L. Li. 2019. Fractional quaternion cosine transform and its application in color image copy-move forgery detection. *Multimed. Tools Appl.* 78, 7 (2019), 8057–8073.
[193] R. Benouini, I. Batioua, K. Zenkouar, A. Zahi, S. Najah, and H. Qjidaa. 2019. Fractional-order orthogonal Chebyshev moments and moment invariants for image representation and pattern recognition. *Pattern Recognit.* 86 (2019), 332–343.
[194] M. Hosny, M. M. Darwish, and T. Aboelenen. 2020. New fractional-order Legendre-Fourier moments for pattern recognition applications. *Pattern Recognit.* 103 (2020), 107324,
[195] H. Zhang, Z. Li, and Y. Liu. 2016. Fractional orthogonal Fourier-Mellin moments for pattern recognition. In *Proceedings of the Chinese Conference on Pattern Recognition*, 766–778.
[196] B. Chen, M. Yu, Q. Su, H. J. Shim, and Y. Q. Shi. 2018. Fractional quaternion Zernike moments for robust color image copy-move forgery detection. *IEEE Access* 6 (2018), 56637–56646.
[197] M. Hosny, M. M. Darwish, and T. Aboelenen. 2020. Novel fractional-order generic Jacobi-Fourier moments for image analysis. *Signal Process.* 172 (2020).
[198] M. Hosny, M. M. Darwish, and T. Aboelenen. 2020. Novel fractional-order polar harmonic transforms for gray-scale and color image analysis. *J. Frankl. Inst.-Eng. Appl. Math* 357, 4 (2020), 2533–2560.
[199] M. Hosny, M. M. Darwish, and M. M. Eltoukhy. 2020. Novel multi-channel fractional-order radial harmonic Fourier moments for color image analysis. *IEEE Access* 8 (2020), 40732–40743.
[200] M. Hosny, M. A. Elaziz, and M. M. Darwish. 2020. Color face recognition using novel fractional-order multi-channel exponent moments. *Neural Comput. Appl.* 33 (2021), 5419–5435.
[201] Y. Qu, C. Cui, S. Chen, and J. Li. 2005. A. fast subpixel edge detection method using Sobel-Zernike moments operator. *Image Vis. Comput.* 23, 1 (2005), 11–17.
[202] A. Iscen, G. Tolias, P. H. Gosselin, and H. Jegou. 2015. A. comparison of dense region detectors for image search and fine-grained classification. *IEEE Trans. Image Process* 24, 8 (2015), 2369–2381.
[203] Z. Chen, S. K. Sun, and A. 2010. Zernike moment phase-based descriptor for local image representation and matching. *IEEE Trans. Image Process.* 19, 1 (2010), 205–219.
[204] M. Schlemmer et al.. 2007. Moment invariants for the analysis of 2D flow fields. *IEEE Trans. Vis. Comput. Graph.* 13, 6 (2007), 1743–1750.
[205] A. Kumar, M. O. Ahmad, and M. N. S. Swamy. 2019. Tchebichef and adaptive steerable-based total variation model for image denoising. *IEEE Trans. Image Process.* 28, 6 (2019), 2921–2935.
[206] A. Kumar. 2017. Deblurring of motion blurred images using histogram of oriented gradients and geometric moments, signal process. *Image Commun.* 55 (2017), 55–65.

[207] X. Gao, Q. Wang, X. Li, D. Tao, and K. Zhang. 2011. Zernike-moment-based image super resolution. *IEEE Trans. Image Process.* 20, 10 (2011), 2738–2747.

[208] Z. Yang and F. S. Cohen. 1999. Cross-weighted moments and affine invariants for image registration and matching. *IEEE Trans. Pattern Anal. Mach. Intell.* 21, 8 (1999), 804–814.

[209] G. Xiao, Y. Lu, W. Zhang, G. Li, and Wang. 2016. *Neurocomputing* 214 (2016), 214, 587–593.

[210] L. Li, W. Lin, X. Wang, G. Yang, K. Bahrami, and A. C. Kot. 2016. No-reference image blur assessment based on discrete orthogonal moments. *IEEE T. Cybern.* 46, 1 (2016), 39–50.

[211] M. Liang, J. Du, L. Li, Z. Xue, X. Wang, F. Kou, and X. Wang. 2020. Video super-resolution reconstruction based on deep learning and spatio-temporal feature self-similarity. *IEEE Trans. Knowl. Data Eng.* (2020). DOI: 10.1109/TKDE.2020.3034261

[212] M. Bronstein, M. M. Bronstein, L. J. Guibas, and M. Ovsjanikov. 2011. Shape Google: Geometric words and expressions for invariant shape retrieval. *ACM Trans. Graph* 30, 1 (2011).

[213] F. Murtagh and J. L. Starck. 2008. Wavelet and curvelet moments for image classification: Application to aggregate mixture grading. *Pattern Recognit. Lett.* 29, 10 (2008), 1557–1564.

[214] G. Paschos, I. Radev, and N. Prabakar. 2003. Image content-based retrieval using chromaticity moments. *IEEE Trans. Knowl. Data Eng.* 15, 5 (2003), 1069–1072.

[215] L. Zhao and L. S. Davis. 2005. Closely coupled object detection and segmentation. *Proc. IEEE Int. Conf. Comput. Vis.* (2005), 454–461.

[216] X. Wang, Z. Wu, L. Chen, H. Zheng, and H. Yang. 2016. Pixel classification based color image segmentation using quaternion exponent moments. *Neural Netw.* 74 (2016), 1–13.

[217] D. Ioannidis, D. Tzovaras, I. G. Damousis, S. Argyropoulos, and K. Moustakas. 2007. Gait recognition using compact feature extraction transforms and depth information. *IEEE Trans. Inf. Forensic Secur.* 2, 3 (2007), 623–630.

[218] N. Vincent and J. M. Ogier. 2019. Shall deep learning be the mandatory future of document analysis problems?. *Pattern Recognit.* 86 (2019), 281–289.

[219] M. Liu and P. T. Yap. 2012. Invariant representation of orientation fields for fingerprint indexing. *Pattern Recognit.* 45, 7 (2012), 2532–2542.

[220] S. M. Lajevardi and Z. M. Hussain. 2010. Higher order orthogonal moments for invariant facial expression recognition. *Digit. Signal Prog.* 20, 6 (2010), 1771–1779.

[221] D. Tsougenis, G. A. Papakostas, D. E. Koulouriotis, and V. D. Tourassis. 2012. Performance evaluation of moment-based watermarking methods: A review. *J. Syst. Softw.* 85, 8 (2012), 1864–1884.

[222] Y. Zhang, X. Luo, Y. Guo, C. Qin, and F. Liu. 2019. Zernike moment-based spatial image steganography resisting scaling attack and statistic detection. *IEEE Access* 7 (2019), 24282–24289.

[223] Y. Zhao, S. Wang, X. Zhang, and H. Yao. 2013. Robust hashing for image authentication using Zernike moments and local features. *IEEE Trans. Inf. Forensic Secur.* 8, 1 (2013), 55–63.

[224] D. Cozzolino, G. Poggi, and L. Verdoliva. 2015. Efficient dense-field copy-move forgery detection. *IEEE Trans. Inf. Forensic Secur.* 10, 11 (2015), 2284–2297.

[225] X. Dai, H. Shu, L. Luo, G. Han, and J. L. Coatrieux. 2010. Reconstruction of tomographic images from limited range projections using discrete radon transform and Tchebichef moments. *Pattern Recognit.* 43, 3 (2010), 1152–1164.

[226] X. Dai and S. Khorram. 1999. A feature-based image registration algorithm using improved chain-code representation combined with invariant moments. *IEEE Trans. Geosci. Remote Sens.* 37, 5 (1999), 2351–2362.

[227] F. Chaumette. 2004. Image moments: A general and useful set of features for visual servoing. *IEEE Trans. Robot. Autom.* 20, 4 (2004), 713–723.

[228] D. Casasent and D. Psaltis. 1977. New optical transforms for pattern recognition. *Proc. IEEE* 65 (1977), 77–84.

[229] L. Zhai, B. Li, J. Chen, X. Wang, M. Xu, J. Liu, and S. Lu. 2018. Chemical image moments and their applications. *Trac-Trends Anal. Chem.* 103 (2018), 119–125.

[230] A. Sit, W. H. Shin, and D. Kihara. 2019. Three-dimensional Krawtchouk descriptors for protein local surface shape comparison. *Pattern Recognit.* 93 (2019), 534–545.

[231] M. Uhrin. 2021. Through the Eyes of a Descriptor: Constructing Complete, Invertible, Descriptions of Atomic Environments. Retrieved April 2021 from https://arxiv.org/abs/2104.09319

[232] J. Lu, V. E. Liong, X. Zhou, and J. Zhou. 2015. Learning compact binary face descriptor for face recognition. *IEEE Trans. Pattern Anal. Mach. Intell.* 37, 10 (2015), 2041–2056.

[233] C. Szegedy et al.. 2015. Going deeper with convolutions. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 1–9.

[234] K. He, X. Zhang, S. Ren, and J. Sun. 2016. Deep residual learning for image recognition. In *Proceedings on the IEEE Conference on Computer Vision and Pattern Recognition*, 770–778.

[235] X. Zhang, C. Liu, and C. Suen. 2020. Towards robust pattern recognition: A review. *Proc. IEEE* 108, 6 (2020), 894–922.

[236] Y. Wang, Y. Y. Tang, L. Li, H. Chen, and J. Pan. 2019. Atomic representation-based classification: Theory, algorithm, and applications. *IEEE Trans. Pattern Anal. Mach. Intell.* 41, 1 (2019), 6–19.

[237] Y. Wang, M. Shi, S. You, and C. Xu. 2016. DCT inspired feature transform for image retrieval and reconstruction. *IEEE Trans. Image Process.* 25, 9 (2016), 4406–4420.

[238] T. Zhao and T. Blu. 2020. The Fourier-Argand representation: An optimal basis of steerable patterns. *IEEE Trans. Image Process.* 29 (2020), 6357–6371.

[239] J, L. Silvan-Cardenas and A. Salazar-Garibay. 1980. Local geometric deformations in the DHT domain with applications. *IEEE Trans. Image Process.* 28 (1980), 4,

[240] A. Sit and D. Kihara. 2014. Comparison of image patches using local moment invariants. *IEEE Trans. Image Process.* 23, 5 (2014), 2369–2379.

[241] W. Tan and A. Kumar. 2014. Accurate iris recognition at a distance using stabilized iris encoding and Zernike moments phase features. *IEEE Trans. Image Process.* 23, 9 (2014), 3962–3974.

[242] A. Kar, S. Pramanik, A. Chakraborty, D. Bhattacharjee, E. S.L. Ho, and H. P.H. Shum. 2021. LMZMPM: Local modified Zernike moment per-unit mass for robust human face recognition. *IEEE Trans. Inf. Forensic Secur.* 16 (2021), 495–509.

[243] Y. Hao, Q. Li, H. Mo, H. Zhang, H. Li, and Ami-Net. 2018. *IEEE Signal Process. Lett.* 25, 7 (2018), 1064–1068.

[244] J. Wu, S. Qiu, Y. Kong, Y. Chen, L. Senhadji, and H. Shu. 2017. MomentsNet: A simple learning-free method for binary image recognition. In *Proceedings of the International Conference on Image Processing*, 2667–2671.

[245] Y. Duan, J. Lu, J. Feng, and J. Zhou. 2017. Learning rotation-invariant local binary descriptor. *IEEE Trans. Image Process.* 26, 8 (2017), 3636–3651.

[246] L. Xie, J. Wang, W. Lin, B. Zhang, and Q. Tian. 2017. Towards reversal-invariant image representation. *Int. J. Comput. Vis.* 123, 2 (2017), 226–250.

[247] L. Sifre and S. Mallat. 2013. Rotation, scaling and deformation invariant scattering for texture discrimination. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 1233–1240.

[248] M. Jaderberg, K. Simonyan, A. Zisserman, and K. Kavukcuoglu. 2015. *Spatial Transformer Networks.* Retrieved June 2015 from https://arxiv.org/abs/1506.02025

[249] Y. Wang, C. Xu, C. Xu, and D. Tao. 2019. Packing convolutional neural networks in the frequency domain. *IEEE Trans. Pattern Anal. Mach. Intell.* 41, 10 (2019), 2495–2510.

[250] A. Agarwal, R. Singh, M. Vatsa, and N. K. Ratha. 2020. Image transformation based defense against adversarial perturbation on deep learning models. *IEEE Trans. Dependable Secur. Comput.* 18, 5 (2020), 2106–2121.

[251] T. Yang, J. Ma, Y. Miao, X. Liu, X. Wang, and Q. Meng. 2019. PLCOM: Privacy-preserving outsourcing computation of Legendre circularly orthogonal moment over encrypted image data. *Inf. Sci.* 505 (2019), 198–214.

Received March 2021; revised July 2021; accepted July 2021