

RESEARCH

Open Access

Image analysis using modified exponent-Fourier moments



Bing He^{1,2}, Jiangtao Cui^{1*}, Bin Xiao³ and Yanguo Peng¹

Abstract

Classic exponent-Fourier moments (EFMs) have been popularly used for image reconstruction and invariant classification. However, EFMs lack natively the translation and scaling-invariant; in addition, they exhibit two types of drawbacks, namely numerical instability and reconstruction error, which in turn influence their reconstruction capability and image classification accuracy. This study considers the challenge of defining modified EFMs (MEFMs), which are based on modified exponent polynomials. In our methods, the basis function of traditional EFMs is appropriately modified, and these modified basis functions are used to replace the original ones. The basis function of the proposed moments is composed of piecewise modified exponent polynomials modulated by a variable parameter exponential envelope. Various types of optimal-order moments can be established by slightly adjusting the bandwidth of the modified basis functions. Finally, we extend the rotation-invariant feature of previous works and propose a new method of scaling and rotation-invariant image recognition using the proposed moments in a log-polar coordinate domain. The translation invariance can then be achieved by an image projection operation, which is substituted for the traditional approach based on the calculation of image geometric moments. The experimental results demonstrate that the MEFMs perform better than traditional EFMs and other classic orthogonal moments including the latest image moments in terms of the image reconstruction capability and the invariant recognition accuracy of smoothing filters, in both noise-free and noisy conditions.

Keywords: Orthogonal moments, Modified exponent-Fourier moments, Image classification, Image reconstruction, Rotation-invariant

1 Introduction

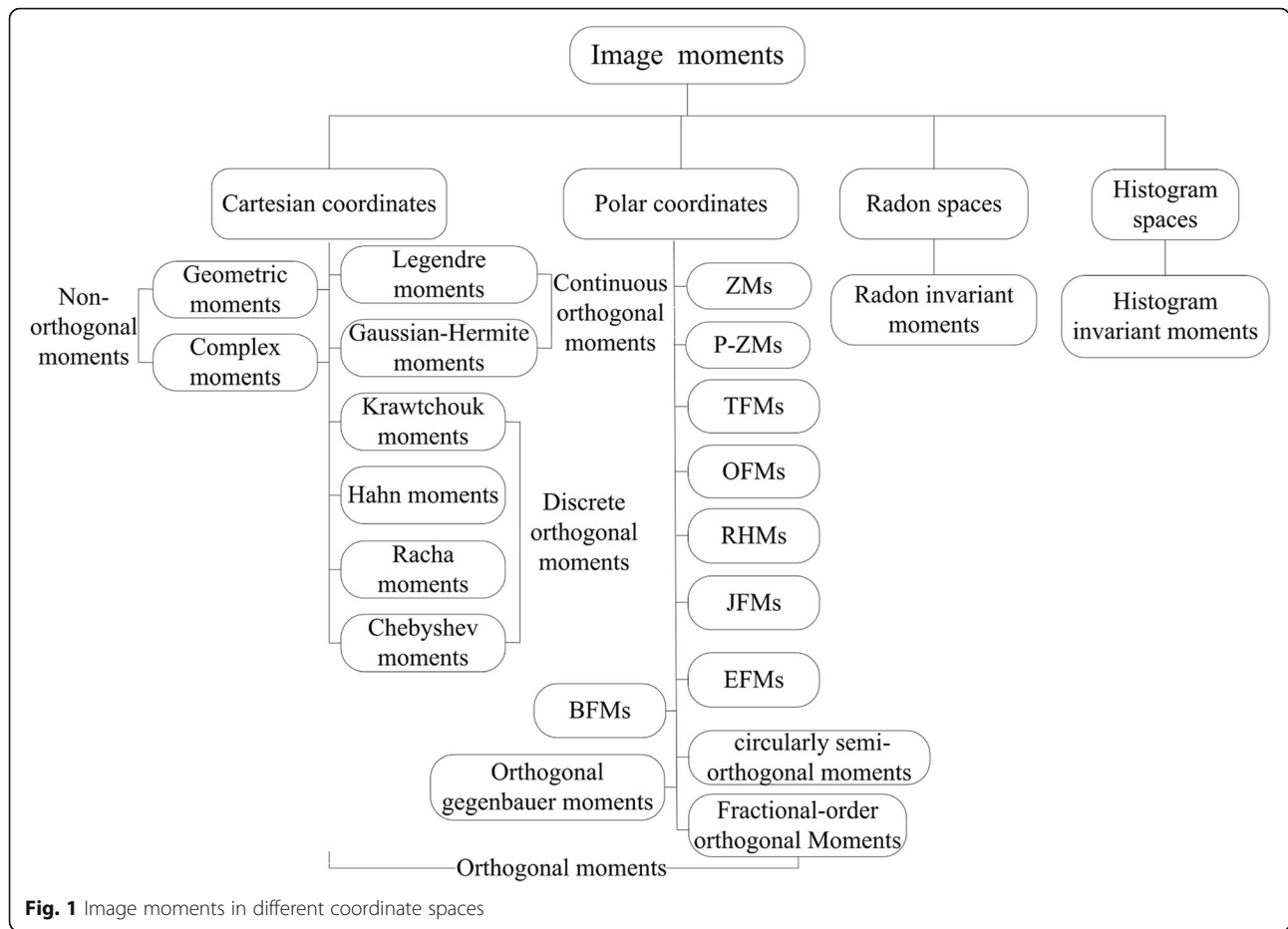
Moments and moment invariants are global descriptors for image feature extraction that have become a hot topic in the field of image analysis. In recent years, various moments have been widely used in image reconstruction [1, 2], image detection [3, 4], target classification [5], digital watermarking [6, 7], image compression [8], and other applications [9, 10]. The study of moments mainly focuses on three directions. The first one is establishing image moments in different coordinate spaces, such as the Cartesian coordinate space [11, 12], polar coordinate space [13, 14], and Radon transformation space [15, 16], among others. The performance of moments reconstructed in the Cartesian coordinate space is better than those in

the polar coordinate space and Radon transformation space. The computation complexity is lower; however, rotation-invariant features are difficult to achieve. The image moments are natively rotation-invariant in the polar coordinate and Radon transformation space, and their geometric invariance can easily be achieved. Therefore, the existing image moments are more greatly established in the polar coordinates. Figure 1 shows the various types of image moments in different coordinate systems. The second direction is studying the description ability of the image moments under different basis functions to search for the best basis functions to construct the image moments with better image reconstruction effect and numerical stability. Generally speaking, traditional image moments do not have the inherent properties of geometric invariance; thus, they need to be restructured and designed to satisfy the geometric invariance in pattern recognition. In summary, the construction of rotation-invariant is

* Correspondence: cuijt@xidian.edu.cn

¹School of Computer Science and Technology, Xidian University, Xi'an 710071, China

Full list of author information is available at the end of the article



becoming a hot topic in the study of image moments, which is the third direction for research of image moments.

As mentioned earlier, the essence of image moments is the set of image transformations based on basis functions. The advantages and disadvantages of its basis functions will directly affect the performance of the constructed image moments. In light of whether the basis set satisfies orthogonal conditions, the image moments can be divided into orthogonal and non-orthogonal moments (similarly known as orthogonal and non-orthogonal transformations, respectively, for example, discrete cosine transform [17], Fourier transform [18], Haar-wavelet transform [19], and Walsh transform [20] belong to orthogonal transformations). Non-orthogonal moments like geometric moments [21], complex moments [22], and rotation moments [23] have made certain achievements in the field of moment applications. The basis functions of non-orthogonal moments are relatively simple with an image reconstruction that is difficult to realize. In addition, the non-orthogonal moments generally have information redundancy that is sensitive to noise. The

orthogonal moments can overcome the disadvantages of the abovementioned non-orthogonal moments, thereby becoming a main focus area in the field of image moments in the recent years.

Orthogonal moments can be defined in different coordinate spaces. The basis functions of orthogonal moments defined in polar coordinates are composed of radial polynomials and Fourier complex exponential factors with angular variables (regarded as amplitude and phase coefficients as well); thus, they are called radial orthogonal moments. The radial orthogonal moments in Fig. 1 mainly include Zernike moments (ZMs) [13], pseudo-Zernike moments (PZMs) [14], orthogonal Fourier–Mellin moments (OFMMs) [6], Jacobi–Fourier moments (JFMs) [24], Tchebichef–Fourier moments (TFMs) [25], radial harmonic-Fourier moments (RHFMs) [26], Bessel–Fourier moments (BFMs) [18, 27], exponent-Fourier moments (EFMs) [7], and radial shifted Legendre moments (RSLMs) [28]. These radial orthogonal moments normally have the basic ability of image reconstruction. Moreover, their significant characteristic is that the radial polynomials satisfy orthogonal condition in the unit circle and natively possess a rotation-invariant feature. Thus,

radial orthogonal moments have become the preferred descriptor for geometric invariant image recognition, especially for rotation-invariant recognition. Basis functions are regular polynomials defined in the Cartesian coordinates, which can be further divided into continuous orthogonal moments and discrete orthogonal moments, such as Legendre moments (LMs) [29] and Gaussian–Hermite moments (GHMs) [2] that belong to continuous orthogonal moments and Tchebichef moments (TMs) [30], Krawtchouk moments (KMs) [31], Hahn moments (HMs) [32], and Racha moments (RMs) [33] that belong to discrete orthogonal moments. Discrete orthogonal moments do not involve any numerical approximation operations; hence, their basis functions can accurately satisfy an orthogonal condition. Consequently, the image reconstruction performance is better than that of traditional continuous orthogonal moments. In addition, we can construct different moments in other spaces like Radon transform invariant moments and histogram invariant moments in the Radon transform space and histogram space, respectively.

Shortcomings still exist in the abovementioned traditional orthogonal moments. On the one hand, the order of the existing orthogonal moments can only be taken as an integer value, which makes the development of orthogonal moments encounter bottlenecks caused by this constraint. To solve this problem, Xiao et al. [34] and Yang et al. [35] proposed fractional orthogonal moments. The integer-order can be extended to a real-order (also known as fractional order) using their proposed models. Further experimental results showed that the fractional order orthogonal moments were better than the traditional orthogonal moments based on the integer order in image reconstruction, noise robustness, and image recognition. Chen et al. [36, 37] recently extended the ZMs and PZMs to a quaternion and a fractional framework for color image feature extraction. The application of image moments has also been further improved. On the other hand, for image sets with larger distinctions, the classification effect is preferable using the lower-order moments constructed using the basis functions of traditional orthogonal moments. However, for the classification effect of the image sets with smaller discrimination, numerical instability will occur when higher-order moments are adopted. The reason is that the basis functions of traditional orthogonal moments are fixed either in lower- or higher-order moments, which can result in poor classification results in pattern recognition. Wang et al. [38] proposed a circularly semi-orthogonal moment that can maintain a good numerical stability in higher-order moments and can obtain a better visual effect in image reconstruction. This method only performs a simple and fixed modulation on the orthogonal basis functions, and the basis functions of

different-order moments are still fixed; hence, the method lacks generality.

Classic orthogonal moments (e.g., EFMs) have the defects of numerical instability and poor accuracy of image recognition in some image classifications, especially in texture image recognition. A modified exponent-Fourier moment (MEFM) is proposed herein based on the concept proposed in [34, 38]. We mainly make attempts in view of three aspects. First, we take on the challenge of studying the performance of semi-orthogonal basis functions at the intersections between the orthogonal and non-orthogonal moments for image reconstruction and pattern recognition. A general semi-orthogonal moment model suitable for different orders can also be established. Second, a new method of the theoretical analysis model of the image moments in the frequency domain is proposed, namely time–frequency correspondence analysis. Finally, a simple and useful algorithm for rotation, scaling, and translation (RST) of invariant image recognition using the proposed moments is introduced herein.

The remainder of this paper is organized as follows: Section 2 provides some preliminaries about the classic exponent-Fourier moments for the 2D images; Section 3 introduces the MEFMs in the polar coordinates and discusses some properties of the MEFMs; Section 4 describes the experiments on the computational complexities of the image moments, image reconstruction, optimal parameter selection, and RST invariant image recognition under both noisy and noise-free, smoothing filter conditions; and Section 5 presents the conclusions.

2 Preliminaries

This section briefly reviews the definition of the classic orthogonal exponent-Fourier moments (EFMs) [39] for an image along with some EFM properties.

2.1 Exponent-Fourier moments

The EFMs of order n with repetition m for a 2D image function $f(r, \theta)$ in the polar coordinates is defined as

$$E_{nm} = \frac{1}{2\pi} \int_0^{2\pi} \int_0^1 f(r, \theta) R_n^*(r) e^{-j m \theta} r dr d\theta \tag{1}$$

where $f(r, \theta)$ denotes the 2D image function in the polar coordinates; $j = \sqrt{-1}$, $n = 0, 1, 2, \dots, \Lambda$, $m = 0 \pm 1, \pm 2, \dots, \Lambda$ represent the moment orders; and $R_n^*(r)$ is the conjugate function of $R_n(r)$ defined as

$$R_n^*(r) = \frac{1}{\sqrt{r}} e^{-j 2 n \pi r} \tag{2}$$

Based on the principle of the orthogonal theory, a 2D image function can be reconstructed by the infinite

series of the orthogonal function $E_{nm}R_n^*(r)$ over the unit circle.

$$\bar{f}(r, \theta) \approx \sum_{n=1}^{n_{\max}} \sum_{m=1}^{m_{\max}} E_{nm}R_n(r)e^{jm\theta} \tag{3}$$

2.2 Properties of EFMs and other radial orthogonal moments

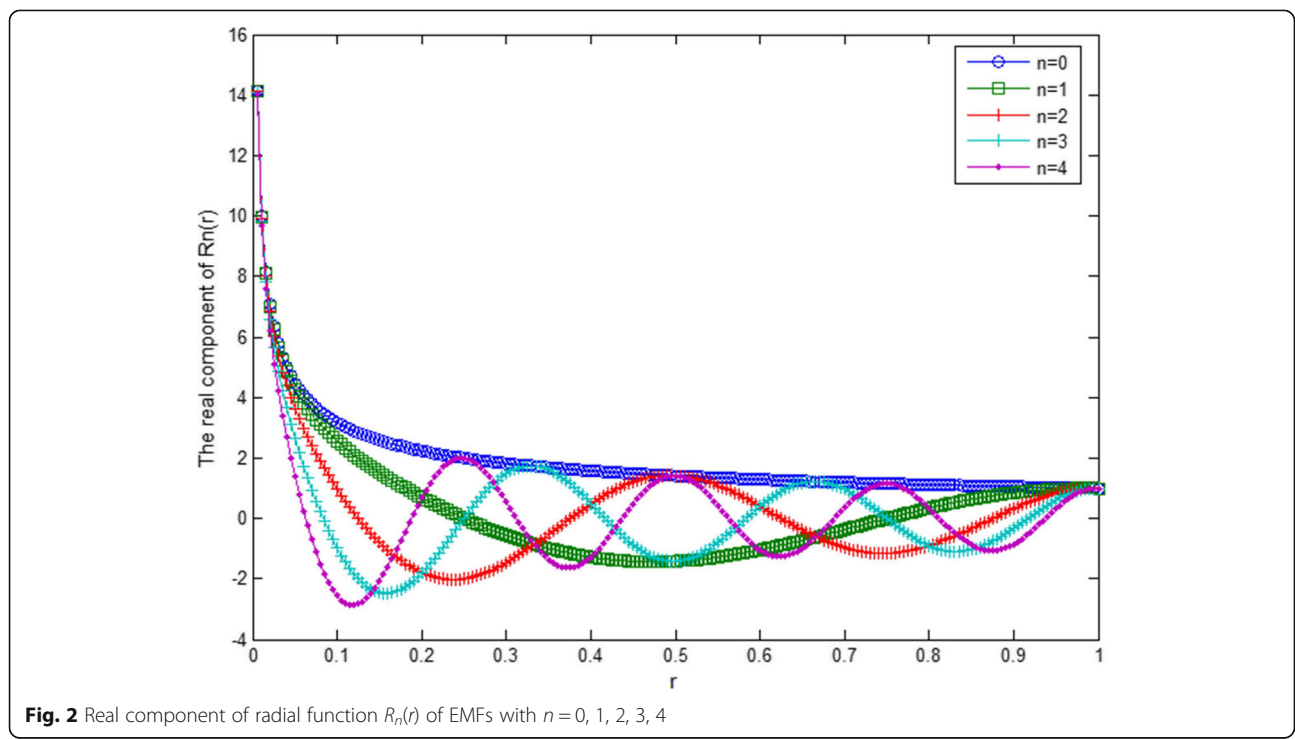
For the existing radial orthogonal moments, the number of zeros of the orthogonal polynomials plays a significant role in describing the high-spatial-frequency components of an image. The real and imaginary parts of the radial polynomial of EFMs have $2n$ and $2n+1$ zeros in the interval $0 \leq r \leq 1$, respectively [39]. Meanwhile, the Bessel polynomials and the orthogonal Fourier–Mellion polynomials have $n+2$ and n zeros in the interval $0 \leq r \leq 1$, respectively [6, 40]. Zernike polynomials only have $(n - m)/2$ zeros in the interval $0 \leq r \leq 1$. Therefore, the degree n of EFMs required to represent an image is much lower than that in BFMs, OFMMs, and ZMs, thereby causing the EFMs to have a stronger capability in describing an image compared to the other orthogonal moments (e.g., BFMs, OFMMs, and ZMs) in the polar coordinates. Additionally, classic EFMs and other radial orthogonal moments have the property of rotation-invariance similar to geometric invariant recognition. The above-mentioned properties show that the exponent-Fourier

moments are potentially useful as feature descriptors for image analysis.

3 Methods

3.1 Analysis of the numerical instability involved in classic EFMs

Hu et al. [39] first proposed classic EFMs based on a radial function $R_n(r)$ shown in Eq. (2), which satisfied the orthogonal condition over interval $0 \leq r \leq 1$. However, the radial function $R_n(r)$ is numerically unstable for classic EFMs, which could cause poor image reconstruction and imprecise image classification in practical applications. The abovementioned reasons are mainly attributed to the following two aspects: First, when r is equal to 0, the real component of the radial function $R_n(r)$ of the EFMs will tend to infinity, and the imaginary part is not a number (i.e., not a number (NaN) value), which are illegal in an actual operation. Second, as shown in Fig. 2, the real component of the radial function $R_n(r)$ of the EFMs will be very large when r tends to 0. This will result in the numerical instability during computation in image moments and will make the computed moments' value inaccurate. Let $r = \Delta r$. When r is equal to 0, where Δr is the minimum value close to 0 (e.g., $\Delta r = 0.005$), the first question can be avoided. However, choosing a suitable value of Δr for the computation in lower- or higher-order moments will be difficult. Furthermore, the second question always exists in the computation of the EFMs all the same.



3.2 Definition of MEFMs

We improve the EFMs and define their modified version, MEFMs, as follows to avoid the numerical instability of the EFMs:

$$M_{nm} = \frac{1}{2} \int_0^{2\pi} \int_0^1 f(r, \theta) T_n(\alpha, r) e^{-jm\theta} r dr d\theta \quad (4)$$

where $f(r, \theta)$ is an image function in the polar coordinates; $n = 0, 1, 2, \dots, \Lambda$, $m = 0, \pm 1, \pm 2, \dots, \Lambda$ are the moments' order; and $T_n(\alpha, r)$ denotes the radial basis functions of the image moments defined as follows:

$$T_n(\alpha, r) = \begin{cases} 16^{-\frac{\alpha_1}{4}r} e^{-j2nmr}, & n \in N_{\text{low}} \\ 16^{-\frac{\alpha_2}{4}r} e^{-j2nmr}, & n \in N_{\text{high}} \end{cases} \quad (5)$$

where $T_n(\alpha_1, \alpha_2, r) = T_n(\alpha, r)$, $(\alpha_1, \alpha_2) \in R$, $n = 0, 1, 2, \dots, N_{\text{low}}$, and N_{high} represents the number of lower- and higher-order moments for the image moments, respectively. The radial basis functions $T_n(\alpha, \beta; r)$ can be comprehended as a set of orthogonal exponent functions $R_n^*(r)$ in Eq. (2) multiplied by the compound envelope factor $\sqrt{r}(16^{-\frac{\alpha}{4}r})$. The basis function $R_n^*(r)e^{-jm\theta}$ is orthogonal over the interior of the unit circle.

$$\int_0^{2\pi} \int_0^1 [R_n^*(r)e^{-jp\theta} r dr d\theta = 2\pi \delta_{nm} \delta_{pq}] \quad (6)$$

where 2π is the normalization coefficient and δ_{nm} or δ_{pq} is the Kronecker delta function. Thus, the MEFMs can also be called semi-orthogonal EFMs.

3.3 Calculation of MEFMs

In the image analysis process, all testing images are digital images; thus, Eq. (4) must be replaced by a discrete form. Consider a digital image $f(x_i, y_j)$ of the $M \times N$ pixels, $0 \leq i \leq M$, $0 \leq j \leq N$. We normalize the $M \times N$ pixels onto the unit circle $[-1, 1] \times [-1, 1]$. Eq. (4) can be rewritten as

$$M_{nm} = \frac{1}{2\pi} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} f(x_i, y_j) \times \int_{x_i-\frac{\Delta x}{2}}^{x_i+\frac{\Delta x}{2}} \int_{y_j-\frac{\Delta y}{2}}^{y_j+\frac{\Delta y}{2}} T_n(\alpha, \sqrt{x^2 + y^2}) e^{-jm(\tan^{-1}(y/x))} dx dy \quad (7)$$

where $x_i = \frac{2i+1-M}{M}$, $y_j = \frac{2j+1-N}{N}$ and $\Delta x = \Delta y = \frac{2}{\sqrt{M^2+N^2}}$

. A zero-order approximation method (ZOA) is used to calculate the double integration in Eq. (7) and make a fair comparison with the classic EFMs in [39] via the following experiments:

$$\int_{x_i-\frac{\Delta x}{2}}^{x_i+\frac{\Delta x}{2}} \int_{y_j-\frac{\Delta y}{2}}^{y_j+\frac{\Delta y}{2}} T_n(\alpha, \sqrt{x^2 + y^2}) e^{-jm(\tan^{-1}(y/x))} dx dy \approx \Delta x \Delta y T_n(\alpha, \sqrt{x_i^2 + y_j^2}) e^{-jm(\tan^{-1}(y_j/x_i))} \quad (8)$$

Substituting Eq. (8) into Eq. (7), the modified EFM can be calculated by ZOA as

$$M_{nm} = \frac{2}{\pi(M^2 + N^2)} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} f(x_i, y_j) T_n(\alpha, \sqrt{x_i^2 + y_j^2}) e^{-jm(\tan^{-1}(y_j/x_i))} \quad (9)$$

Similarly, the reconstructed image can be expressed by the following formula:

$$f(x_i, y_j) \approx \sum_{n=1}^{b_{\text{max}}} \sum_{m=1}^{m_{\text{max}}} \tilde{M}_{nm} T_n(\alpha, \sqrt{x_i^2 + y_j^2}) e^{jm(\tan^{-1}(y_j/x_i))} \quad (10)$$

3.4 Computation complexity and stability analysis of MEFMs

All the computations of image moments, including the other moments used for comparison, are implemented by the ZOA algorithm proposed in [38] to fairly compare and efficiently verify the properties of the MEFMs without considering accurate calculation and fast algorithm of image moments. Compared with the existing classic orthogonal moments based on higher-order polynomials (e.g., ZMs, LMs, OFMMs, and BFMs), the proposed radial polynomial of the MEFMs is simple (i.e., it is only composed of trigonometric and exponential functions with parameter variables). In practice, it does not involve factorial and accumulative summation operations in classic orthogonal moments; thus, the computational complexity is lower. The radial polynomial of ZMs, OFMMs, and BFMs in lower-order moments ($n = 10$) [40] is basically close to the uniform distributions, and the amplitudes are more stable (e.g., the amplitude of the radial polynomials of ZMs and BFMs is located in the interval of $[-1, 1]$, while there are only a few lower-order moments of OFMMs, whose amplitudes exceed 2, and the rest are located in the interval of $[-2, 2]$). However, with the increase in the order of the image moments, a numerical instability will appear in the calculation of the abovementioned classic orthogonal moments (e.g., Fig. 3 shows the numerical distribution curves of the classic orthogonal moments at higher-order moments, order $n = 50$). Figure 3 shows that the radial polynomial of ZMs and OFMMs is close to 0 in

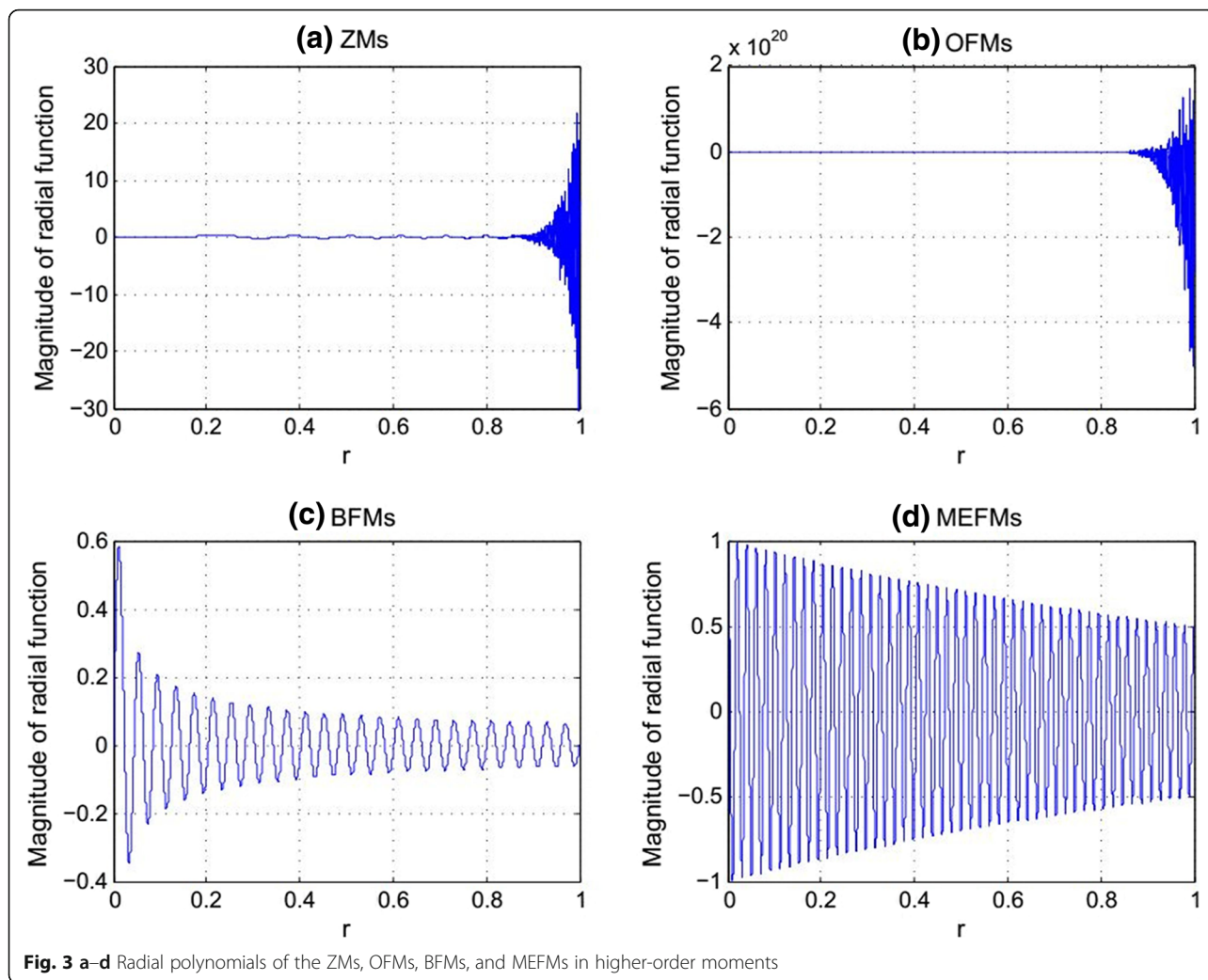


Fig. 3 a-d Radial polynomials of the ZMs, OFMs, BFMs, and MEFMs in higher-order moments

the interval $[0, 0.8]$. Each amplitude gradually increases in the interval of $(0.8, 1)$, and the numerical values tend to be unstable (i.e., the amplitude of OFMMs is close to 1.7×10^{20} , when $r = 0.95$). The radial polynomial of BFMs also tends to decay in the interval $[0, 1]$ (e.g., the amplitude is attenuated to $[-0.1, 0.1]$ in the interval of $[0.5, 1]$). However, the polynomial of MEFMs is almost uniform when the order $n = 50$, and the amplitude is stable. In addition, the classic EFMs [39] and RHFMs [26] have introduced factor $\sqrt{1/r}$ into their radial basis functions to satisfy orthogonal condition; however, the polynomials' amplitude of the EFMs and RHFMs tend to NaN (non-number) and Inf (infinity), respectively, when $r = 0$. This will result in a numerical instability in the image moments constructed. Compared with the other classic orthogonal moments, the proposed image moments can avoid this phenomenon and make the constructed moments more stable. The orthogonal moments constructed by the orthogonal

polynomials are better than the non-orthogonal moments in terms of the overall performance. However, this does not mean that the orthogonal polynomials are in a stable state at each point in the defined domain; thus, the proper correction of its unstable orthogonal basis functions can make the image moments reach their best performance. This is the major purpose of the proposed image moments in this paper.

3.5 Time-frequency analysis of MEFMs

From the time-domain point of view, the ZOA theory can effectively explain the properties of the constructed basis functions of the orthogonal moments (i.e., the location of zeros of the radial function and the number of zeros of the radial function represent the sampling position and the sampling frequency of an image, respectively). The higher the number of zeros and the more even the distribution in a region, the better is the

reconstructed image. For a given order n and repetition m , the radial polynomial of BFM and OFMMs has $n + 2$ and n zeros in the interior of a unit circle, respectively, while the radial polynomial of the ZMs only has $(n - m)/2$ zeros in the interval $0 \leq r \leq 1$ [40]. Among radial polynomials (or radial functions) with trigonometric functions as basis functions, the real and imaginary parts of the radial polynomial of EFMs [39] and polar harmonic Fourier moments (PHT) [26] have $2n$ and $2n + 1$ zeros in the interior of the unit circle, respectively. The radial polynomials of the polar sine transforms (PST) [41], polar cosine transforms (PCT) [41], and circularly semi-orthogonal moments (SOMs) similarly have $n + 2$ zeros. Meanwhile, the real and imaginary parts of the radial polynomial of MEFMs have $2n$ and $2n + 1$ zeros in the interior of the unit circle, respectively. As illustrated in Fig. 4, the curve distribution of the real part of the radial polynomial of the MEFMs is smoother at the lower-order moments. This is then compared with the classic orthogonal moments (i.e., ZMs, OFMs, and BFM) and other orthogonal moments with trigonometric functions as basis functions (i.e., EFMs, PHT, and PCT), which are closer to the uniform distribution and whose magnitudes are more stable (e.g., the amplitude distribution interval is $[-1, 1]$). For image recognition, most of the algorithms use the lower-order moments of the image moments as the feature extraction for classification. The lower-order moments have a good robustness to noise in pattern recognition; however, the orders of image moments should be increased to obtain more image feature points as the classification features and make a more precise classification for the image sets under a higher similarity condition (e.g., texture images). Therefore, we need to deeply study the higher-order moments of the image moments. The lower-order moments generally correspond to the low-frequency components of an image (e.g., contours or shapes of an image), while the higher-order moments of the image moments represent the detail components of an image (i.e., high-frequency components). The method of the time-domain analysis can be used for the quantitative analysis of the lower-order moments of the image moments, but it cannot provide a more reasonable description of the high-frequency components of an image (corresponding higher-order moments) for image processing or analysis. In view of the above-mentioned reasons, a method of time–frequency correspondence is proposed from the frequency domain perspective. This method can analyze and improve the stability of different order moments for image recognition. The basic objective is to consider the representation of the basis functions of the image moments in the frequency domain as a 2D filter. We hope that the frequency bandwidth corresponding to

the basis functions of the image moments is wider at the lower-order moments, and the attenuation of the cut-off frequency is as slow as possible. While the corresponding bandwidth is narrower in the higher-order moments, and the attenuation of the frequency cutoff is as fast as possible, in this study, a parameter-modulated MEFM is still proposed and used to verify our concept (Fig. 5). The main principle is to change the bandwidth in the frequency domain by adjusting parameter α of the radial function of the MEFMs in the time domain. In the low-frequency region of the image (lower-order moments), we want to change parameter α (e.g., $\alpha = 2$ in the experiments) to make the bandwidth as wide as possible, such that more image features of the lower-order moments can be obtained. In the high-frequency region (higher-order moments), the bandwidth is made as narrow as possible by changing parameter α (e.g., $\alpha = 0.2$ in the experiments), such that more high-frequency components can be suppressed, especially noise interference. Finally, the theoretical results are illustrated and verified by the experimental results of image reconstruction (Section 4.2).

4 Results and discussion

In this section, the experimental results are used to validate the theoretical framework developed in the previous sections. This section includes four subsections. In the first subsection, we discuss the computational complexities of MEFMs as compared to those of BFM, ZMs, OFMs, PST, and PCT. In the second subsection, the question of how well an image can be represented using MEFMs is addressed, and the image reconstruction capability of MEFMs is compared with those of BFM, ZMs, OFMs, SOMs, PST, and PCT. In the third subsection, the question of optimal parameter selection for image reconstruction and recognition is discussed. A new method for the RST invariant image recognition using the proposed moments and the experimental study on the RST recognition accuracy of MEFMs is provided in the last subsection.

4.1 Computational complexities

In this section, we demonstrate in terms of the computation time exactly how less complex the computation of the radial polynomial of the MEFMs is when compared to those of BFM, ZMs, and OFMs. Table 1 shows a summary of the comparisons of the computation time for computing the radial polynomials between MEFMs and the other radial orthogonal moments. In the calculation, the order of the image moments is 5, 10, 15, ..., 30. The test image is a Lena gray-level image

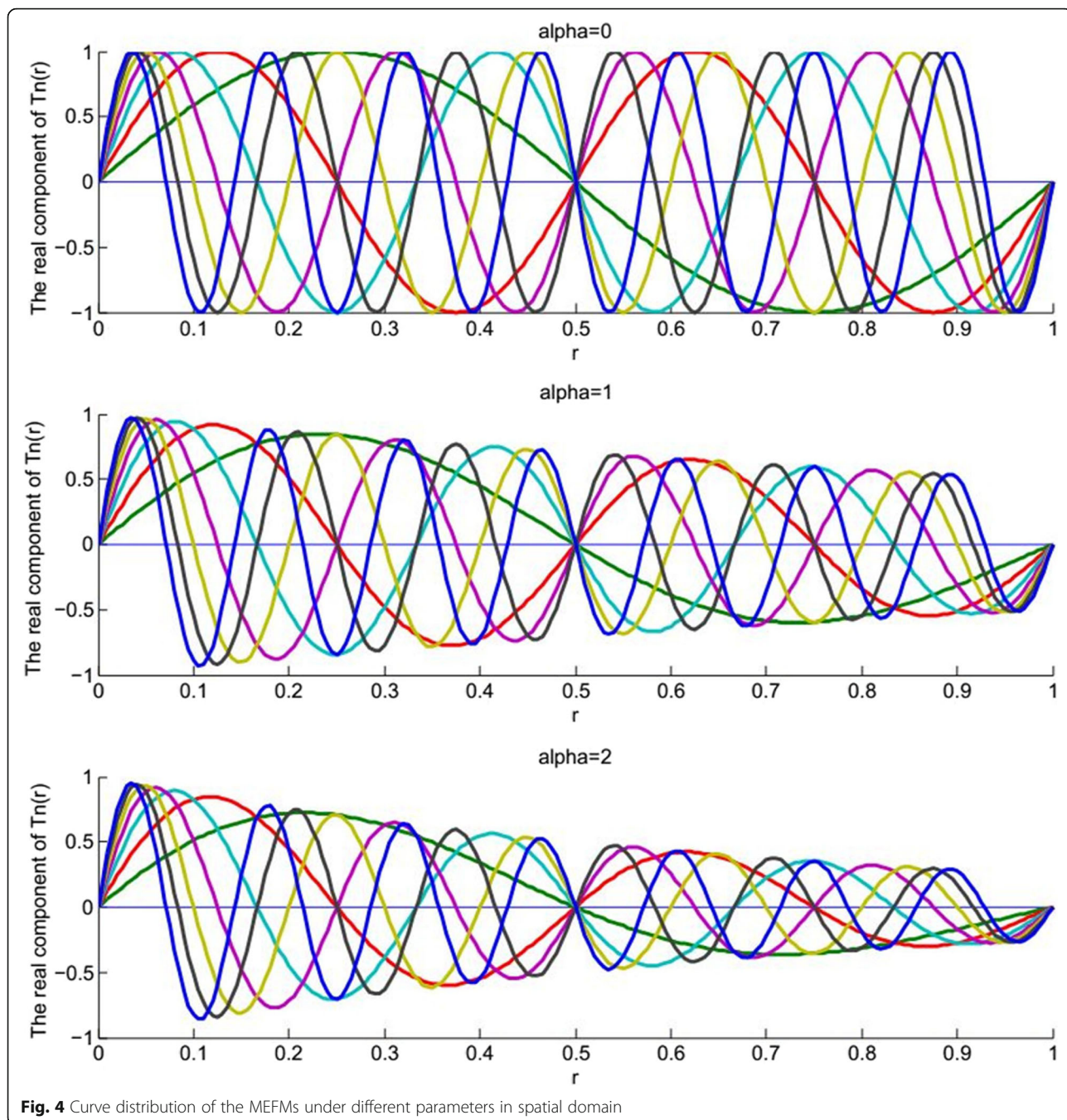


Fig. 4 Curve distribution of the MEFMs under different parameters in spatial domain

(Fig. 6), while the size is 128×128 . The average value of the computation time by six different order moments is taken as the time-consuming measurements for all the image moments. The hardware configuration of the test computer comprises a 3.2 GHz Intel^(R) Core (TM) i5 CPU and 8 GB memory. The software is MATLAB R2013a. Table 1 shows that the time consumed by the MEFMs is slightly higher than that of the PST and EFMs, but its computing time is significantly lower

than that of the other classical orthogonal moments (i.e., ZMs, OFMs, and BFM).

4.2 Image reconstruction

In this subsection, the image representation capability of the MEFMs is presented. For the convenience of computing the image moments, the number of moments used in the image reconstruction and recognition experiments is limited based on $n_{\max} = m_{\max} = N$,

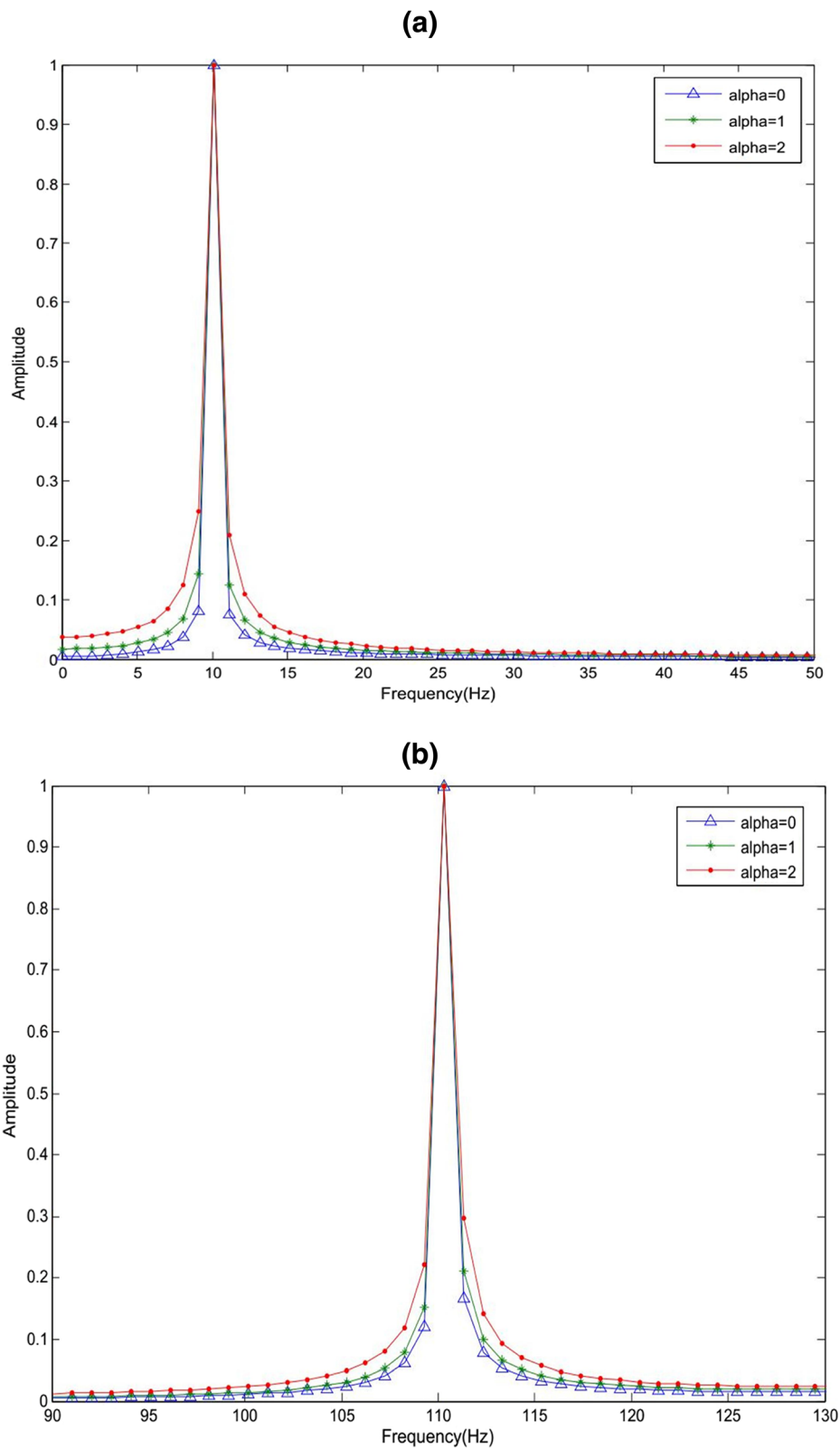


Fig. 5 Characteristic curves of the MEFMs under different parameters in frequency domain. **a** In lower-orders moments. **b** In higher-orders moments

Table 1 Basis functions computation time

Number	Image moments	Computation time (ms)
1	ZMs	113.9
2	OFMs	144.8
3	BFMs	158.3
4	EFMs	37.2
5	PST	30.6
6	PCT	42.9
7	MEFMs	40.3

$N \in Z^+$. In addition, we use the statistical-normalization image reconstruction error (SNIRE) defined in [34] to measure the performance of the image reconstruction.

$$\bar{\epsilon}^2 = \frac{\sum_{x=1}^N \sum_{y=1}^N |f(x,y) - \bar{f}(x,y)|^2}{\sum_{x=1}^N \sum_{y=1}^N f^2(x,y)} \tag{11}$$

Where $f(x,y)$ is the original image and $\bar{f}(x,y)$ is the reconstructed image.

4.2.0.1 Experiment 1 A set of binary images including digits from 0 to 9 and the uppercase English letters from A to J, and another set of gray-level images and color images including Lena, cameraman, woman and plane, baboon, and pepper, as shown in Fig. 7 are used as test images. The size of each image is 64×64 . The proposed MEFMs are obtained from the



Fig. 6 Lena gray-level image

images shown in Fig. 7, and the images are reconstructed using the maximum order of 35 and the parameter α of the radial function of the MEFMs is 2. The results are given in Fig. 8. It can be seen from Fig. 8 that by using the proposed MEFMs, either color, gray-level, or binary images can be reconstructed well.

4.2.0.2 Experiment 2 To demonstrate the validity of the theory related to the proposed method of time-frequency correspondence in Section 3.5. A comparison of the proposed moments for image reconstruction ability in different parameters is performed and a binary image of uppercase English letter E, a gray-level image cameraman, and a color image baboon are considered in the experiment. The reconstructed experimental results from two types of different methods of determining parameters (i.e., $\alpha = 0$ and $\alpha = 2$) in lower-order moments (e.g., the order $N = 5, 7, 9, 11,$ and 13) and higher-order moments (e.g., the order $N = 55, 60, 65, 70, \dots, 120$) are shown in Figs. 9, 10, and 11. Incidentally, lower-order moments and higher-order moments of image moments are related to image reconstruction, e.g., let order N of image moments be 10 and 100, respectively. $N = 10$ is considered to be a lower-order moment, while $N = 100$ is a higher-order moment. The comparison study of the reconstructed images using the MEFMs in two types of different parameters ($\alpha = 0$ and $\alpha = 2$) shows that, in lower-order moments ($N = 5, 7, 9, 11,$ and 13), the subjective vision of the reconstructed images under parameter $\alpha = 2$ is better than the reconstructed image when $\alpha = 0$, the objective evaluation standard related to the performance of image reconstruction has illustrated this phenomenon as well, i.e., the SNIRE of parameter $\alpha = 2$ in the lower-order moments is generally less than that of parameter $\alpha = 0$. However, with the increase of a moment's orders, when the order N of the image moments exceeds 15, the performance for image reconstruction of MEFMs is just opposite to that of lower-order moments. As shown in Table 2, the SNIRE of parameter $\alpha = 0$ in the higher-order moments (e.g., the order N exceeds 25) is less than parameter $\alpha = 2$, when the moments' order $N = 65$, the SNIRE reaches the minimum, and the reconstructed binary image of uppercase English letter "E" is almost close to the original image. As can be seen from Figs. 9 and 10, the subjective vision of reconstructed gray-level and color images under parameter $\alpha = 0$ is better than that the reconstructed under parameter $\alpha = 2$ in higher-order moments. The above experimental results also verify the reliability and rationality of the proposed method with respect to time-frequency correspondence in Section 3.5, i.e.,

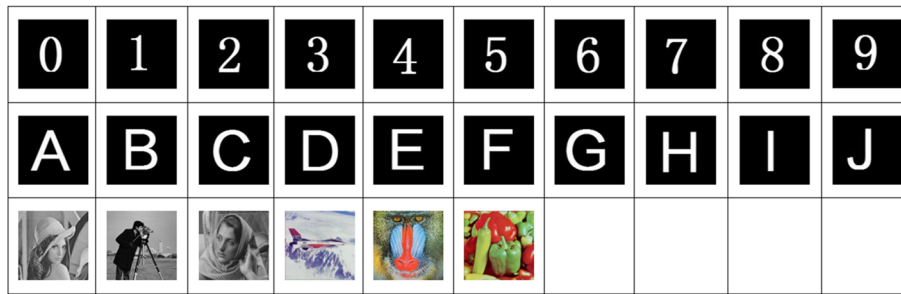


Fig. 7 Binary images, gray-level images, and color images used as test images, each of size 64 × 64

the radial function (or polynomial) of the proposed image moments (MEFMs), whose bandwidths and cutoff frequencies in frequency domain will affect the quality and numerical stability of image reconstruction. If lower-order moments are used to describe the image features, the bandwidth of the radial polynomial of the MEFMs can be adjusted to be slightly larger (e.g., the parameter $\alpha = 2$), while to obtain more image features (the reconstructed images using higher-order moments), the adjustment of bandwidth for radial polynomial of the MEFMs are as narrow as possible (e.g., the parameter $\alpha = 0$).

4.2.0.3 Experiment 3 According to the characteristic analysis of the MEFMs’ radial function in frequency domain, we propose a method of image projection transformation for an original image using piecewise function (or polynomial), when an image is reconstructed at lower-moments and higher-moments, respectively (see Eq. (5) in Section 3.2). In order to verify the validity of the piecewise function in Eq. (5), the proposed image moments (MEFMs) are compared with the ZMs, OFMMs, BFMs, and EFMs in this study, and simulation experiments are performed by the reconstruction of the binary image of uppercase English letter “E.” From the experimental results of

Fig. 11, it is known that the performance of the proposed image moments constructed by the basis functions, which consists of piecewise polynomials in Section 3.2 is superior to other classical orthogonal moments either in lower-order moments or higher-order moments. Especially with the increase in the order of moments, and when the order is $N = 40$, the reconstructed images by OFMMs is invalid. When the order is $N = 50$, the reconstructed images using ZMs is invalid, and the reconstructed images using BFMs and EFMs can maintain good numerical stability in higher-order moments, but those visual effect of image reconstruction are obviously worse than that of the proposed image moments (MEFMs).

We choose the image moments (e.g., SOMs, PST, and PCT) with trigonometric functions as the radial basis functions to reconstruct images and compare the results with the MEFMs to further verify the validity of the proposed image moments. The experimental results show that the SNIRE of the MEFMs along with SOMs, PCT, and PST approximately linearly decreases with the increase in the moments’ order at lower order moments. Moreover, the quality of the reconstructed images is gradually improved. The curve of Fig. 12a shows that the image reconstruction ability of the proposed image moments is

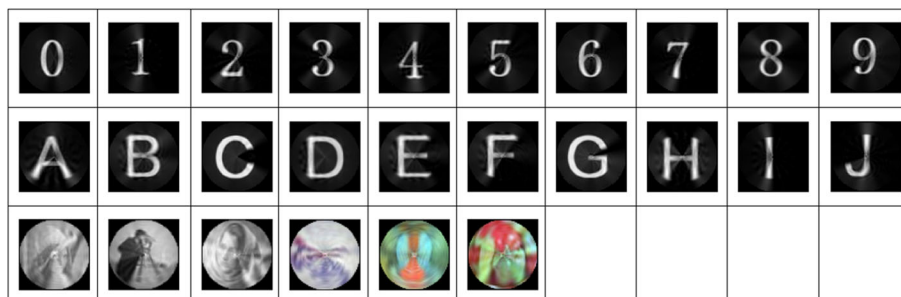
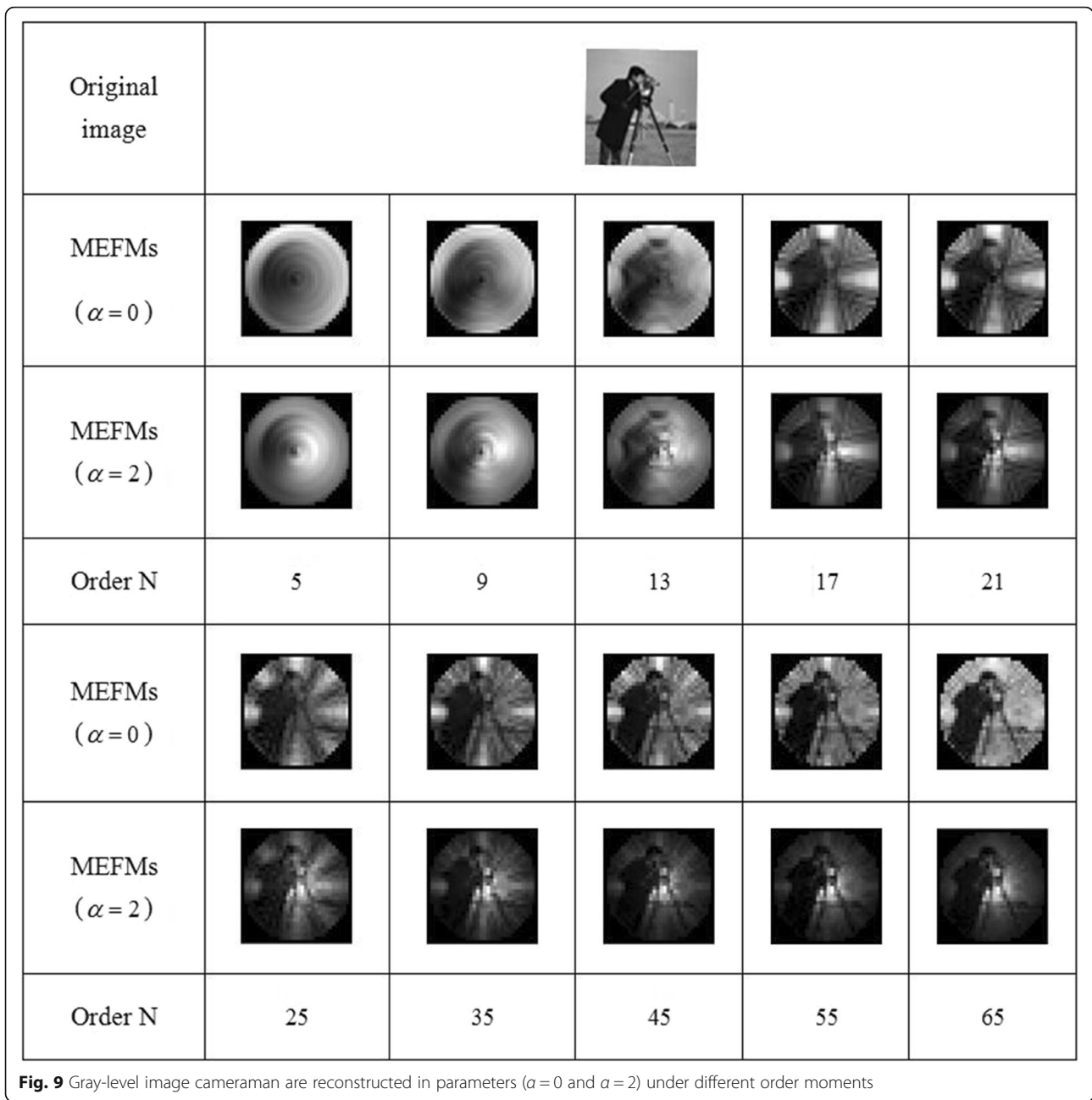


Fig. 8 Images reconstructed from the proposed modified exponent-Fourier moments up to order 35

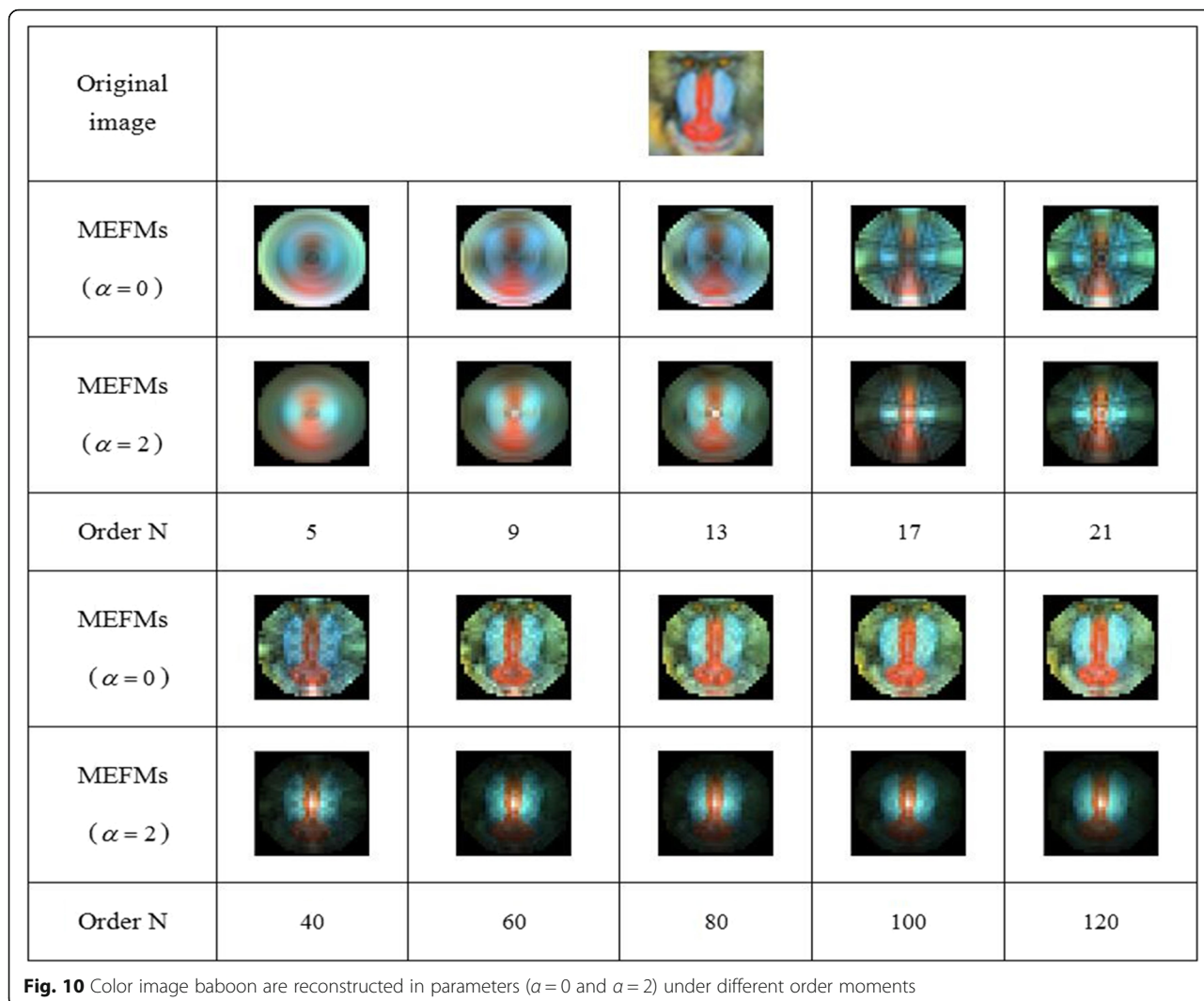


better than that of the EFMs, PCTs, and PSTs. However, with the increase of the moments' order in higher-order moments, Fig. 12b illustrates that the SNIRE of the image moments with trigonometric functions as the radial basis functions does not linearly decrease, and numerical instability exists during image reconstruction. On the contrary, the proposed image moments (MEFMs) can keep the SNIRE gradually decreasing with the increase of moments' order, showing that the performance of image reconstruction in higher-order moments is better than that

of image moments with trigonometric functions as the radial basis functions.

4.3 Optimal parameter selection for image reconstruction and recognition

Based on the analysis theory of the time–frequency correspondence in Section 3.5, the selection of parameter value α in Eq. (5) is crucial for the proposed image moments (MEFMs) that will affect the image reconstruction accuracy and the image recognition



rate. In other words, choosing the optimal parameter value α to obtain a better image description ability is a problem that needs to be solved at the present. Therefore, a selection method of parameter α must be selected for the proposed MEFMs, which could lead to desirable results in image reconstruction. The selection of parameter α is equivalent to an unconstrained optimization problem (i.e., $\min\{\bar{\varepsilon}^2[f, \bar{f}; \alpha, N]\}$) if two variables α and N are limited based on $\alpha_{\min} \leq \alpha \leq \alpha_{\max}$, $N_{\min} \leq N \leq N_{\max}$. For the unconstrained optimization problems, the genetic algorithm (GA) is the most popular and effective method in the recent years. Using GA computing in the proposed image moments, more precise values of parameters α and N can be obtained. However, considering the complexity of the GA implementation process, a simpler algorithm is adopted herein to realize the optimization of parameter α . If the order N of the proposed image moments is fixed, the unconstrained

optimization problem of double variables is transformed into the unconstrained optimization problem of a single variable. The specific implementation process is presented below.

First, we will employ 20 gray-level images selected from the Coil-20 database [42] presented in Fig. 16 and use $D_g(\alpha)$ to evaluate the best selection of parameter α defined as follows to investigate the influence of parameter α on the performance of our introduced method:

$$D_g(\alpha) = \frac{1}{g} \sum_{n=1}^g \bar{\varepsilon}^2[f_n, \bar{f}_n; \alpha] \tag{12}$$

where $g = 20$ denotes the number of gray-level images from the Coil-20 database, and f_n and \bar{f}_n represent the n th original and reconstructed images, respectively. A lower value of D_g indicates a better performance of the proposed image moments in image reconstruction or recognition.

(a)

Original image							
ZMs							
OFMs							
BFM							
EFM							
MEFM ($\alpha = 0$)							
Order N	40	45	50	55	60	65	70

(b)

Original image					
ZMs					
OFMs					
BFMs					
EFMs					
MEFM ($\alpha = 2$)					
Order N	3	8	12	16	21

Fig. 11 a, b Uppercase English letter "E" are reconstructed in parameters ($\alpha = 0$ and $\alpha = 2$), under different moments (MEFMs, ZMs, OFMs, and BFM)

Table 2 Uppercase English letter “E” are reconstructed in parameters ($\alpha = 0$ and $\alpha = 2$) under different order moments

Order N	SNIRE	
	MEFMs($\alpha = 0$)	MEFMs($\alpha = 2$)
5	0.5419	0.5197
7	0.4752	0.4062
9	0.4247	0.3860
11	0.3682	0.3168
13	0.3380	0.3368
15	0.3330	0.3647
17	0.3609	0.4173
19	0.3478	0.3847
21	0.3488	0.3924
25	0.2900	0.4871
30	0.2856	0.4849
35	0.2808	0.5094
40	0.1760	0.5080
45	0.1076	0.5271
50	0.0638	0.5160
55	0.0374	0.5276
60	0.0311	0.5269
65	0.0244	0.5251

Let us consider herein the influence of orthogonality on the basis function of the proposed image moments (e.g., $T_n(\alpha, r) = e^{-j} 2n\pi r$, which is orthogonal in the interval $[0,1]$ when $\alpha = 0$). Note that the search interval is restricted to $-\frac{7}{2} \leq \alpha \leq \frac{7}{2}$ (i.e., we empirically take a value close to zero, and the stepping increment is 0.5) in the experiments. While the order N of the proposed image moments is given ($N = 10$ in lower-order moments and $N = 60$ in higher-order moments in experiments), some different values of D_g can be obtained in terms of the corresponding parameter value α summed up in Table 3. Table 3 clearly shows that $\{N = 10, \alpha = 2\}$ is optimal in lower-order moments, and $\{N = 60, \alpha = 0\}$ is optimal in higher-order moments, which are more appropriate for the task of image reconstruction or recognition. Finally, we can conclude that this experiment could considerably help in selecting the optimal parameter value α for the image reconstruction and classification tasks in the future. The optimal value of parameter α given in Table 3 is also consistent with the conclusion of the time–frequency correspondence method proposed in Section 3.5.

4.4 Rotation, scaling, and translation invariant image recognition

In this section, a new RST invariant system for MEFMs that can be implemented in two steps is proposed: for

translation invariance, the proposed image projection approach can be considered as a new alternative of the traditional algorithm (i.e., the method for the image translation invariant based on calculating the image geometric moments [28] and center moments [43]), followed by extending the basis functions of the MEFMs from the polar coordinate space to the log-polar space, such that the MEFMs have invariant properties of scaling and rotation at the same time.

4.4.1 Scaling and rotation invariance of MEFMs

4.4.1.1 Log-polar mapping In the image processing and recognition process, the original image is usually acquired in a Cartesian coordinate system. First, let $f^{sr}(x, y)$ be the scaled and rotated image of an image function $f(x, y)$ with the scaling factor σ^{-1} and the rotation angle ϕ in the Cartesian coordinates. We then have

$$f^{sr}(x, y) = f(\sigma^{-1}(x \cos\phi + y \sin\phi), \sigma^{-1}(y \cos\phi - x \sin\phi)) \tag{13}$$

Using the conversion relationship from the Cartesian coordinate system to the log-polar coordinate space: $x = e^\rho \cos \theta, y = e^\rho \sin \theta, 0 \leq \theta \leq 2\pi, \rho \in \mathcal{R}^2$, we can rewrite Eq. (13) as

$$\begin{aligned} f^{sr}(e^\rho \cos\theta, e^\rho \sin\theta) &= f(\sigma^{-1}(e^\rho \cos\theta \cos\phi + e^\rho \sin\theta \sin\phi), \\ &\quad \sigma^{-1}(e^\rho \sin\theta \cos\phi - e^\rho \cos\theta \sin\phi)) \\ &= f(\sigma^{-1}e^\rho \cos(\theta-\phi), \sigma^{-1}e^\rho \sin(\theta-\phi)) \\ &= f(e^{-\ln\sigma} e^\rho \cos(\theta-\phi), e^{-\ln\sigma} e^\rho \sin(\theta-\phi)) \\ &= f(e^{\rho-\ln\sigma} \cos(\theta-\phi), e^{\rho-\ln\sigma} \sin(\theta-\phi)) \end{aligned} \tag{14}$$

The above equation can be simply expressed as

$$f^{sr}(\rho, \theta) = f(\rho - \ln\sigma, \theta - \phi) \tag{15}$$

The Fourier transform (FT) of a 2D image function $f(\rho, \theta)$ in the log-polar coordinates can be denoted as follows:

$$f^{sr}(\rho, \theta) \leftrightarrow F(u, v) \tag{16}$$

According to the translation characteristic of 2D Fourier transform, for $f^{sr}(\rho, \theta)$ we have

$$f(\rho - \ln\sigma, \theta - \phi) \leftrightarrow F(u, v) e^{-2\pi j(u \ln\sigma + v\phi)} \tag{17}$$

Thus, it is straightforward that $|F(u, v) e^{-2\pi j(u \ln\sigma + v\phi)}| = |F(u, v)|$.

The above equations and Fig. 13 show that the geometric transformation of the image scaled and rotated in the Cartesian coordinate system will be converted into the corresponding translation operation in the log-polar

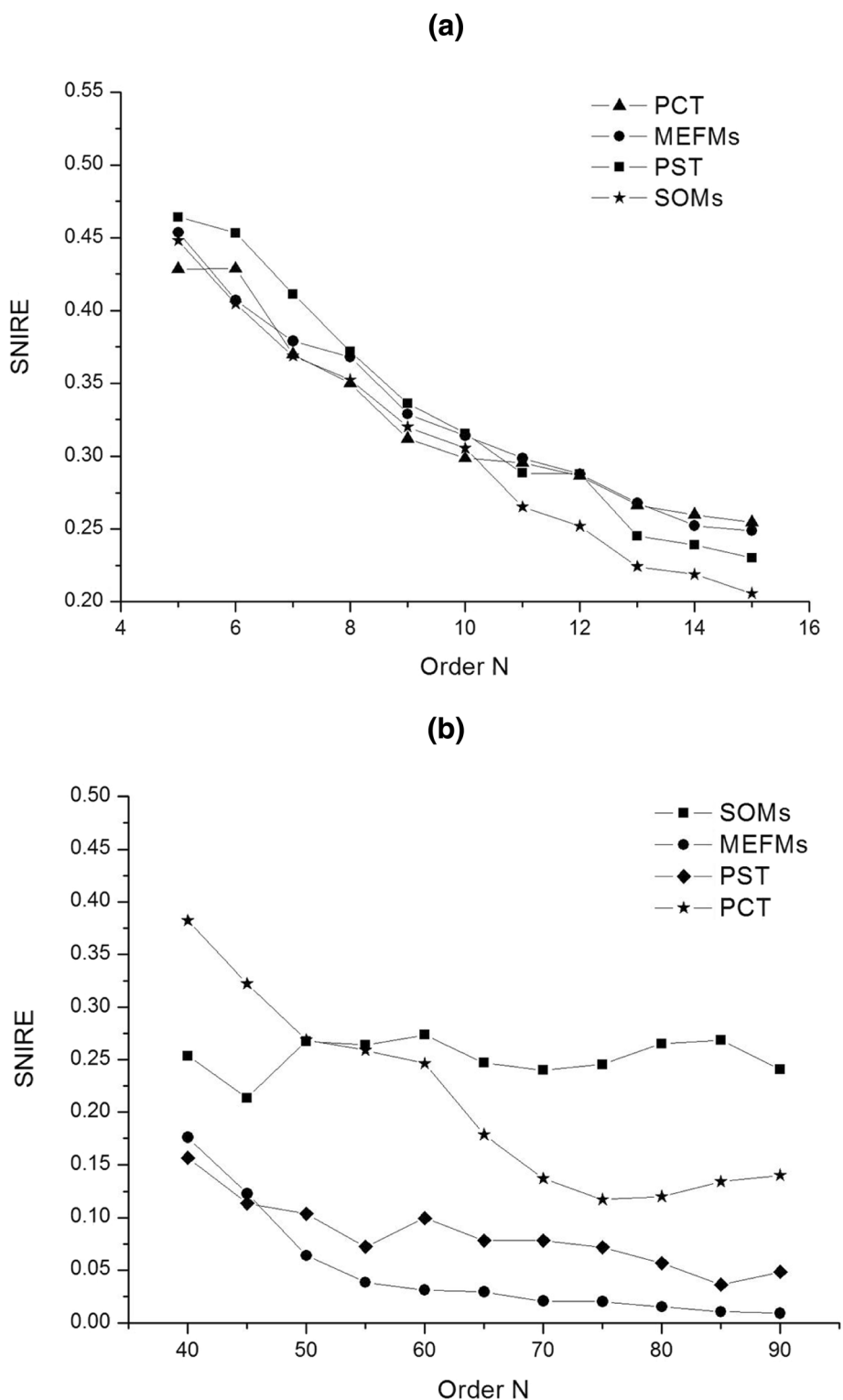


Fig. 12 Comparison of the results of the reconstructed images under MEFMs, SOMs, PCT, and PST. **a** In lower-orders moments. **b** In higher-orders moments

Table 3 The search results of the D_g according to different parameter value α

D_g		
α	Order $N = 10$	Order $N = 60$
-3.5	0.6272	0.6458
-3	0.5841	0.6012
-2.5	0.5344	0.5283
-2	0.4733	0.4942
-1.5	0.4137	0.4090
-1	0.3462	0.3126
-0.5	0.2842	0.2081
0	0.2243	0.1431
0.5	0.1778	0.1535
1	0.1600	0.1942
1.5	0.1514	0.2575
2	0.1421	0.3302
2.5	0.2011	0.4567
3	0.2443	0.5649
3.5	0.3057	0.6529

coordinate space, followed by 2D Fourier transform for $f^r(\rho, \theta)$; thus, the invariance of image scaling and rotation can be achieved.

4.4.1.2 MEFMs invariant computing method in the log-polar coordinate space Encouraged by the success of the Log-polar mapping (LPM) approach and some related works in [44], we take on the challenge of extending the basis functions of MEFMs from the polar coordinates to the log-polar coordinate space, such that the scaling and rotation invariance for the proposed MEFMs can be easily achieved. In light of Eq. (5), we have

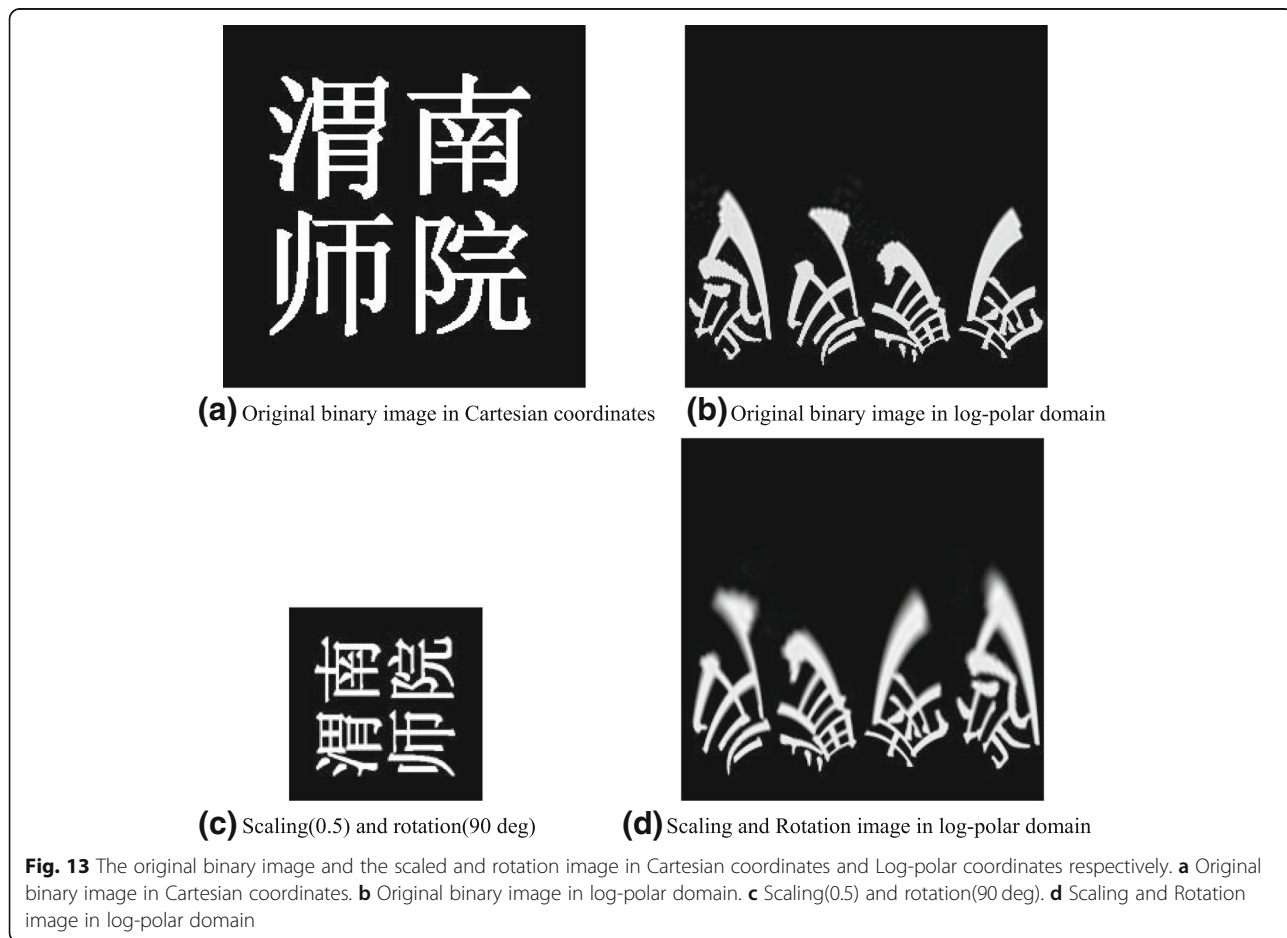
$$T_n(\alpha, r) = |T_n(\alpha, r)|e^{-j2n\pi r} \tag{18}$$

We then let

$$g(r, \theta) = w(\alpha, r)f(r, \theta) \tag{19}$$

where $w(\alpha, r) = |T_n(\alpha, r)|r$ can be regarded as a weighted function, and $g(r, \theta)$ is a weighted image in the polar coordinate system.

Thus, Eq. (4) can be rewritten as follows:



$$\begin{aligned}
 M_{nm} &= \frac{1}{2\pi} \int_0^{2\pi} \int_0^1 f(r, \theta) T_n(\alpha, r) e^{-jm\theta} r dr d\theta \\
 &= \frac{1}{2\pi} \int_0^{2\pi} \int_0^1 f(r, \theta) |T_n(\alpha, r)| e^{-j2n\pi r} e^{-jm\theta} r dr d\theta \\
 &= \frac{1}{2\pi} \int_0^{2\pi} \int_0^1 g(r, \theta) e^{-j2n\pi r} e^{-jm\theta} dr d\theta
 \end{aligned} \tag{20}$$

Similarly, by using the conversion relationship from the polar coordinate system to the log-polar coordinate space: $\rho = \ln r$, $\theta = \theta$, $0 \leq \theta \leq 2\pi$, $\rho \in (-\infty, 0]$, we can change the above Eq. (20) in the polar coordinate domain to the log-polar domain. The modified version of the radial basis function of MEFMs is redefined as

$$T_n(\rho) = e^{j2n\pi\rho} \tag{21}$$

which satisfies the orthogonal condition:

$$\int_{-1}^0 T_n(\rho) T_m^*(\rho) d\rho = \delta_{nm} \tag{22}$$

Hence, the basis function of MEFMs in the log-polar domain can be represented as

$$P_{nm}(\rho, \theta) = T_n(\rho) e^{jm\theta} \tag{23}$$

P_{nm} also satisfies the following orthogonal condition:

$$\int_0^{2\pi} \int_{-1}^0 P_{nm}(\rho, \theta) P_{lk}^*(\rho, \theta) d\rho d\theta = \delta_{nl} \delta_{mk} \tag{24}$$

In light of these conclusions, the modified version of the MEFMs in the log-polar domain is defined as

$$M_{nm}^{LPM} = \int_0^{2\pi} \int_{-1}^0 g(\rho, \theta) P_{nm}^*(\rho, \theta) d\rho d\theta \tag{25}$$

Let $g^{sr}(\rho, \theta)$ denote the scaled and rotated version of an image $g(\rho, \theta)$ with the scaling Factor σ and rotation angle ϕ in the log-polar coordinates. We then have

$$\begin{aligned}
 g^{sr}(\rho, \theta) &= g(\ln(\sigma r), \theta + \phi) \\
 &= g(\ln \sigma + \ln r, \theta + \phi) \\
 &= g(\rho + \ln \sigma, \theta + \phi)
 \end{aligned} \tag{26}$$

Thus, according to Eqs. (24) and (25), the MEFMs of $g^{sr}(\rho, \theta)$ are

$$\begin{aligned}
 \tilde{M}_{nm}^{LPM} &= \int_0^2 \int_{-1}^0 g^{sr}(\rho, \theta) P_{nm}^*(\rho, \theta) d\rho d\theta \\
 &= \int_0^{2\pi} \int_{-1}^0 g^{sr}(\rho, \theta) e^{-j2n\pi\rho} e^{-jm\theta} d\rho d\theta \\
 &= \int_0^{2\pi} \int_{-1}^0 g(\rho + \ln \sigma, \theta + \phi) e^{-j2n\pi\rho} e^{-jm\theta} d\rho d\theta
 \end{aligned} \tag{27}$$

Let $\hat{\rho} = \rho + \ln \sigma$ and $\hat{\theta} = \theta + \phi$, we then have $\rho = \hat{\rho} - \ln \sigma$ and $\theta = \hat{\theta} - \phi$. Eq. (27) can be rewritten as

$$\begin{aligned}
 \tilde{M}_{nm}^{LPM} &= \int_0^{2\pi} \int_{-1}^0 g(\hat{\rho}, \hat{\theta}) e^{-j2n\pi(\hat{\rho} - \ln \sigma)} e^{-jm(\hat{\theta} - \phi)} d\hat{\rho} d\hat{\theta} \\
 &= \int_0^{2\pi} \int_{-1}^0 [g(\hat{\rho}, \hat{\theta}) e^{-j2n\pi\hat{\rho}} e^{-jm\hat{\theta}} d\hat{\rho} d\hat{\theta}] e^{j2n\pi \ln \sigma} e^{jm\phi} \\
 &= \int_0^{2\pi} \int_{-1}^0 [g(\hat{\rho}, \hat{\theta}) e^{-j2n\pi\hat{\rho}} e^{-jm\hat{\theta}} d\hat{\rho} d\hat{\theta}] e^{j2n\pi \ln \sigma} e^{jm\phi} \\
 &\quad M_{nm}^{LPM} e^{j2n\pi \ln \sigma + jm\phi}
 \end{aligned} \tag{28}$$

and we have

$$|\tilde{M}_{nm}^{LPM}| = |M_{nm}^{LPM} e^{j2n\pi \ln \sigma + jm\phi}| = |M_{nm}^{LPM}| \tag{29}$$

Equations (28) and (29) show that the scaling and rotation of an image by a scaling factor of σ and an angle of ϕ result in a shift of the MEFMs in the ρ -axis and θ -axis, respectively. This simple property leads to the conclusion that the magnitudes of the MEFMs of the scaled and rotated image function remain identical to those before scaling and rotation. Thus, the magnitudes $|M_{nm}^{LPM}|$ of the MEFMs can be taken as a scaling and rotation invariant feature for image recognition. For the discretization calculation for scaling and rotation invariance of MEFMs, see [Appendix](#).

P_{nm} is a complete orthogonal basis function; hence, a 2D image can be reconstructed by M_{nm}^{LPM} and represented by the following formula:

$$\begin{aligned}
 g(\rho, \theta) &= \sum_{n=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} M_{nm}^{LPM} P_{nm}(\rho, \theta) \\
 &= \sum_{n=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} M_{nm}^{LPM} e^{j2n\pi\rho} e^{jm\theta}
 \end{aligned} \tag{30}$$

If we maintain the constraints $n \leq n_{\max}$ and $m \leq m_{\max}$, an approximate version of the 2D image function denoted as $\tilde{g}(\rho, \theta)$ can be calculated as

$$\begin{aligned} \tilde{g}(\rho, \theta) &= \sum_{n=1}^{n_{\max}} \sum_{m=1}^{m_{\max}} M_{nm}^{LPM} P_{nm}(\rho, \theta) \\ &= \sum_{n=1}^{n_{\max}} \sum_{m=1}^{m_{\max}} M_{nm}^{LPM} e^{j2n\pi\rho} e^{jm\theta} \end{aligned} \tag{31}$$

4.4.2 Projection approach for the image translation invariance

The existing moments-based image translation invariance was mainly achieved by calculating the image geometric moments or center moments [28, 43]. Its major drawbacks include being time-consuming and a more complex computation process in image recognition (Fig. 14a). The main reason is that if other orthogonal moments (e.g., ZMs, OFMs, and BFM) are used to extract the image features in the image recognition process, the geometric moments or center moments must be calculated again to achieve the translation invariance. Considering the shortcomings of the existing methods, this study proposes a new translation invariant algorithm, also known as the image projection-based method. Our basic approach is to treat translation invariance to separate the target image from the background image. The approach for image translation can be summarized as follows (for the chief algorithm process, see Fig. 14b):

- (1) If the original image is a color image, the color image should first be gray-scale; otherwise, this step can be a default.
- (2) Otsu’s algorithm [45] is used to determine the thresholds of the gray-scale image in the global region and then binarize the gray-scale image according to the thresholds.
- (3) A high-quality binary image can be obtained via a simple image pre-processing operation for binary images (e.g., denoising, filtering, etc.).
- (4) Calculating the projection image in the horizontal direction of the binary image and obtaining the position for the troughs of the projection image, segmentation is performed for the whole image according to the trough point.
- (5) The projection operation in the vertical direction is the same as that in Step 4.
- (6) Finally, according to the segmentation position of the binary image, the target image can be separated from the background in the original image. The experiments are performed on the selected cartoon cat color images (Fig. 15) from the Columbia University image library database [42]. Figure 15 shows (a) as the process of using the projection approach for the untranslated images and (b) as the process of using the projection approach for the translated images.

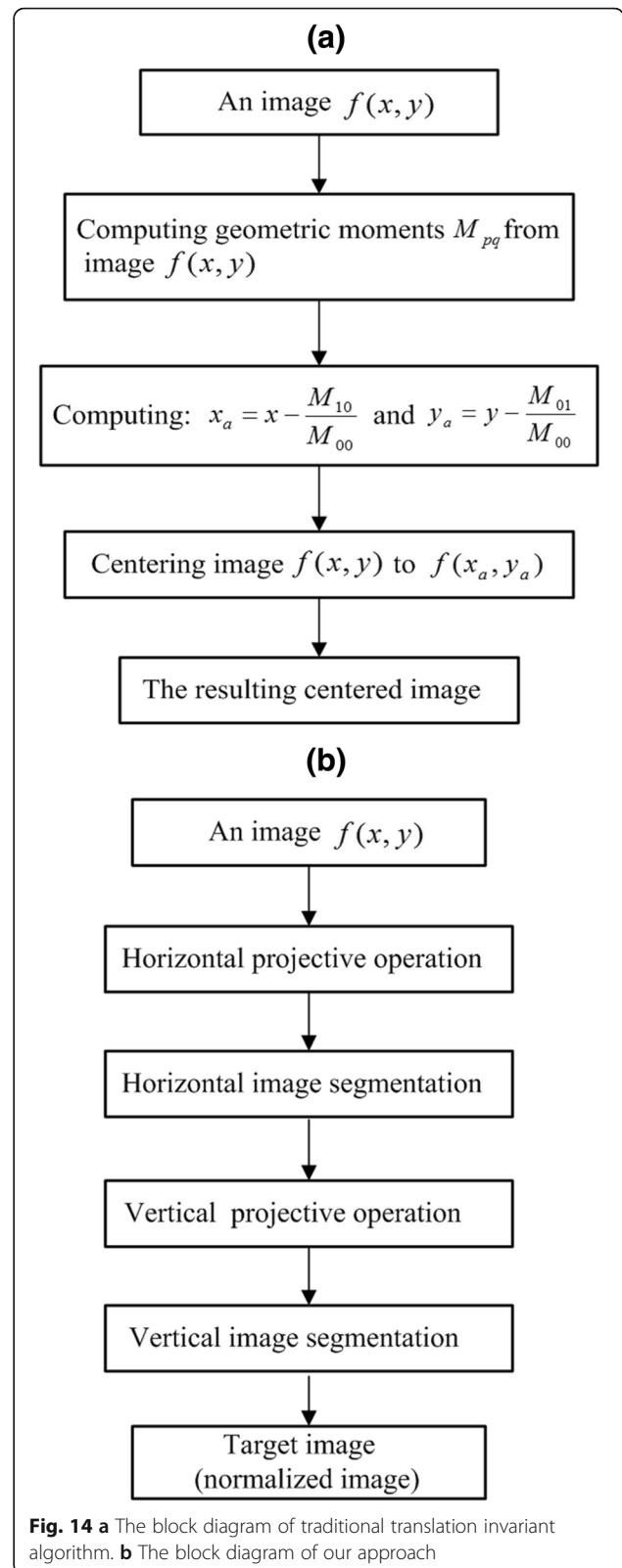
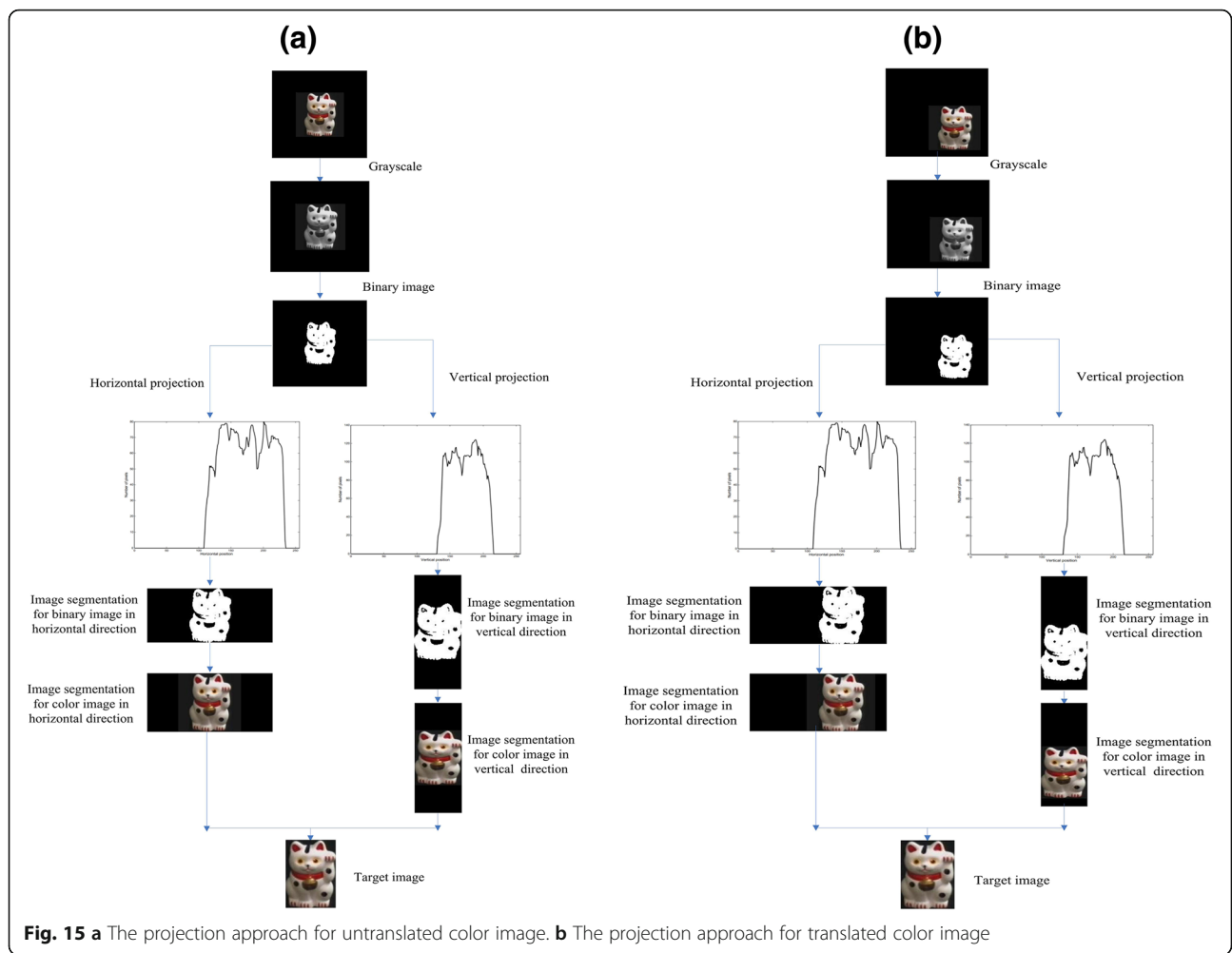


Fig. 14 a The block diagram of traditional translation invariant algorithm. **b** The block diagram of our approach



4.4.3 Test of classification results for the RST invariance

This subsection presents a simulation experimental study on the RST invariant image classification accuracy of the MEFMs under both noisy and noise-free conditions and in a smoothing filter. A comparison with the accuracy of the classic radial orthogonal moments (e.g., ZMs, OFMs, BFMs, and EFMs) and certain latest orthogonal image moments (e.g., fractional order Legendre moments (Fr-LMs) [2017] and SOMs [2016]) is also depicted. Accordingly, 10th lower-order moments and 60th higher-order moments were adopted and the magnitudes of these selected image moments were used as features for the image classification task. A *k*-nearest neighbor classifier was used to execute the classification. To evaluate the performance of the classification results, the expression of the correct classification percentages (CCPs) is introduced as follows:

$$CCPs = \frac{\text{Number of correctly classified objects}}{\text{The total number of classified objects}} \times 100 \tag{32}$$

4.4.3.1 Datasets

The image classification performance of the proposed methods was evaluated with three test datasets: D1, D2, and D3 (Fig. 16). D1 was produced by selecting pictures from a publicly available database, named Coil-100, from Columbia University (the size of each image was 128 × 128; see [42]) (Fig. 17). D2, including 20 butterfly images, was collected from the internet. Some of these images are shown in Fig. 18 and available in [46]. The Brodatz texture image database was used for D3, which included 112 texture images (the size of each image was 640 × 640). Figure 19 shows the typical 35 pictures in the D3 dataset.

4.4.3.2 Experiment 1

First, the training set, including 200 (100 × 2) images, was constructed by rotating each image in the D1 dataset through angles of 0 and 180°. In the next step, each image of the training set was arbitrarily translated with $(\Delta x, \Delta y) \in [-45, 45]$, subsequently letting $\phi_i = 5 * i$ be a rotation angle vector with *i* being an integer and varying from 0 to 35, rotated by ϕ_i , and scaled with a scaling factor of $\alpha = 0.5 + (2.5 * \theta_i) /$



Fig. 16 Samples of Coil-20 grayscale images from Columbia University



Fig. 17 Some typical color images of Coil-100 from Columbia University



Fig. 18 Some of butterfly color images used in the image classification

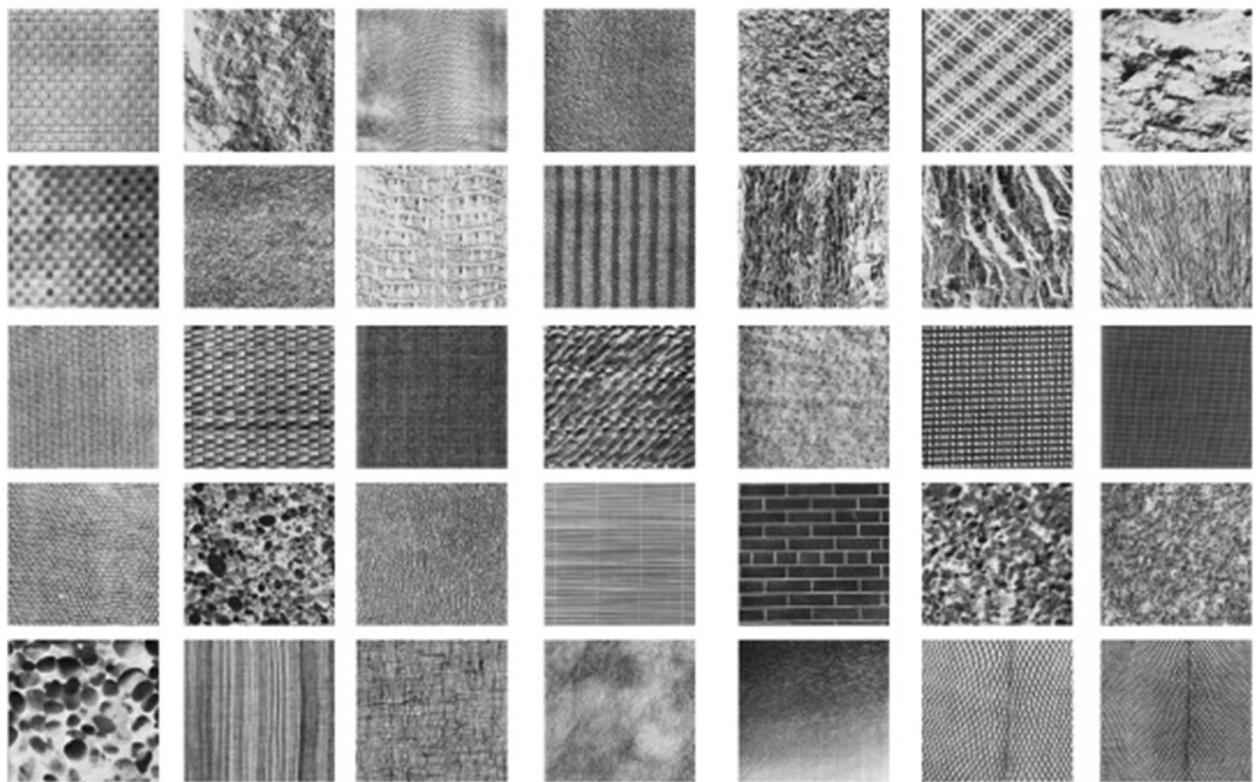


Fig. 19 Some sample images in Brodatz image database

$360 \in [0.5, 3]$. Therefore, we obtained a testing set, including 7200 (36×200) images in the experiment. The proposed approach and the ZM-, OFM-, BFM-, and EFM-based methods were also used to implement classification. In the last step, each of the testing set mentioned earlier was corrupted with the salt and pepper noise with the noise densities varying from 5% to 20% with 5% increment steps. The CCPs were obtained via our method and via the ZM-, OFM-, and BFM-based methods, with the results shown in Table 4. Either in the lower-order moments (10th order moments) or higher-order moments (60th order moments), the image classification performance of our approach (MEFMs) performed better than the other classic orthogonal moments-based methods.

4.4.3.3 Experiment 2 Twenty color butterfly images in dataset D2 with a size of 640×480 were used as the training set. Testing sets, including 1440 (72×20) images, were achieved in the same manner as Experiment 1. The correct CCPs were obtained via our proposed method (MEFMs) and the latest image moments (e.g., Fr-LM-, SOM-, and EFM-based methods). Table 5 shows the summarized classification results. For the texture image classification, the proposed image moments (MEFMs) had superior rotation and scaling invariance. Under the condition of simultaneous rotation and scaling attacks, the proposed MEFMs still maintained a higher classification accuracy compared to the latest image moments in recent years (i.e., Fr-LMs, SOMs, and EFMs).

4.4.3.4 Experiment 3 We further evaluated the image classification capability of the proposed method (texture recognition task). In this experiment, the training set was formed in the same manner as Experiment 1 by rotating each image in dataset D3 through angles 0, 45, 90, and 180°. We can obtain 448 (112×4) images for the database of the training set. The testing set was composed to mix the RST effect with translation $(\Delta x, \Delta y) \in [-45, 45]$ and scaling factor $\alpha \in [0.5, 3]$, which includes 4032 (36×112) images. For the testing set, the CCPs were obtained via our proposed method (MEFMs), EFMs, Fr-LMs, and SOMs. Each image of the testing set

Table 5 Comparison our approach results with the published results in recent years for D2 dataset

The number of moments	CCPs			
	EFMs [2014]	Fr-LMs [2017]	SOMs [2016]	Ours [proposed]
5	68.7%	75.5%	76.4%	78.2%
10	71.8%	80.2%	82.2%	82.5%
60	89.3%	95.5%	97.3%	97.7%
80	91.6%	98.9%	98.4%	99.3%

was divided into two groups: the first group was corrupted by salt and pepper noise with noise densities varying from 5% to 20% with steps of 5% increments, while the second was manipulated by a smoothing filter with different smooth windows (Fig. 20 and Table 6). All these image moments' CCPs inclined to reduce as the density of salt and pepper noise and smooth windows increased; however, the reductions of the proposed MEFMs were the least in these mentioned methods. This result proves that of the four types of image moments, our proposed image moments (MEFMs) exhibited the highest robustness for the salt and pepper noise as well as the smoothing filter operations.

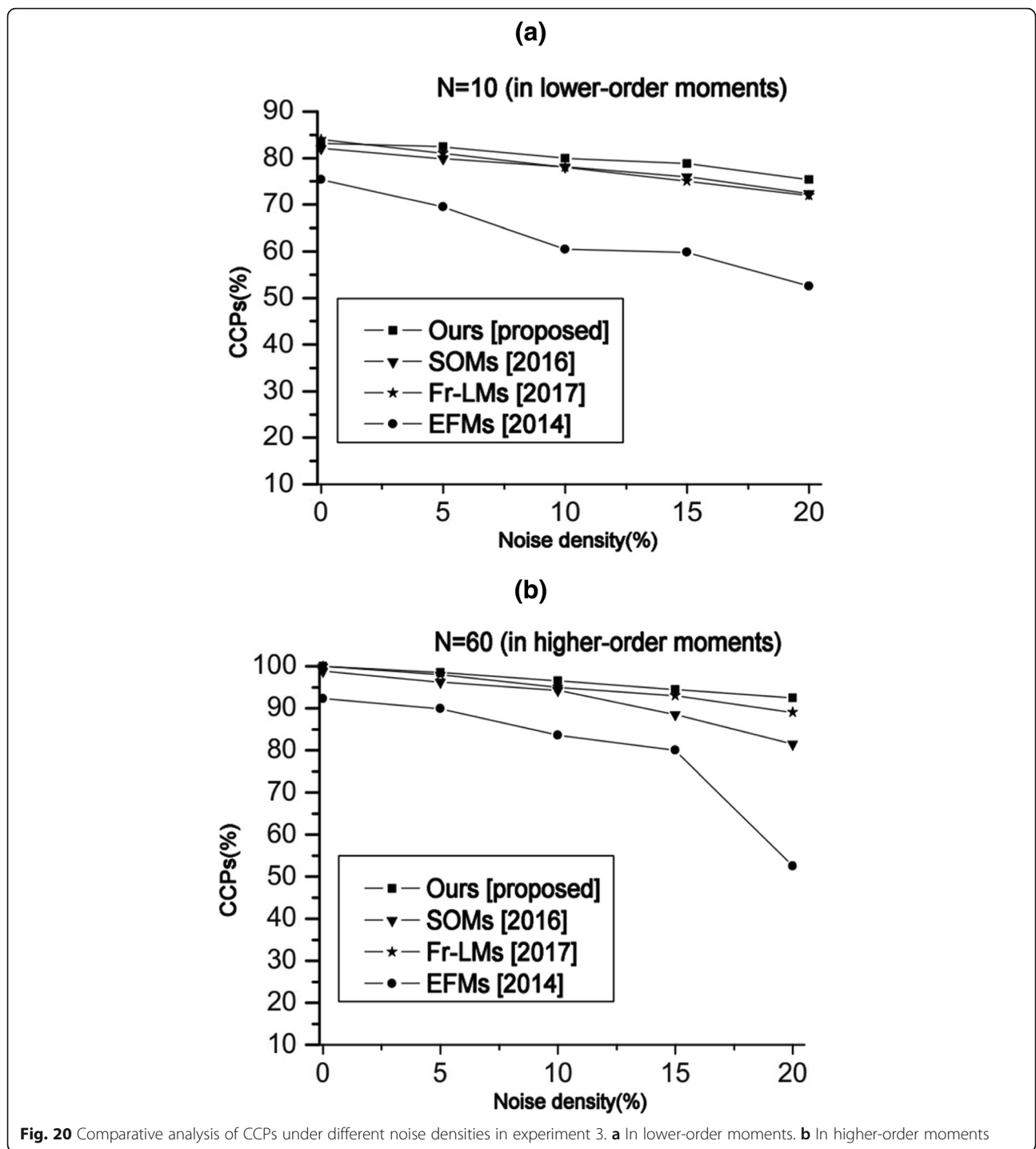
5 Conclusions

This study introduced a new set of moments based on the modified exponent polynomials, called MEFMs. The main contributions of this study are as follows:

- (1) A new type of piecewise modified exponent polynomial, also known as the semi-orthogonal polynomial, was derived. The derived polynomial is the transformed versions of classical exponent polynomial.
- (2) To build a series of numerically stable different-order image moments for image reconstruction and pattern recognition, a new method of time–frequency correspondence is proposed herein, which can improve the image reconstruction effect and accuracy of image recognition.
- (3) We propose a new method for RST invariant recognition and compared it with the traditional

Table 4 Comparative study of the CCPs by various methods (including classic ZMs, OFMs, BFMs, and ours) for D1 dataset contaminated by salt and pepper noise

CCPs (%)	KNN							
	$N = 10$ (lower-order moments)				$N = 60$ (higher-order moments)			
	ZMs	OFMs	BFMs	Ours	ZMs	OFMs	BFMs	Ours
Noise densities								
0%	65.8	71.6	72.2	71.9	91.7	98.8	99.3	100
5%	61.6	68.5	70.6	71.2	90.2	95.8	94.9	99.8
10%	58.8	62.9	67.2	68.1	88.6	90.6	89.8	95.3
15%	55.7	60.3	62.8	64.7	86.3	89.8	89.4	94.6
20%	50.3	57.2	60.7	60.9	81.5	85.5	86.1	91.4



moment invariant-based method. Our approach is more practical and effective for geometric invariant recognition. For the future work, we will search for superior semi-orthogonal image moments for local feature extraction in the image analysis because non-moment-based methods (e.g., [47, 48]) can effectively extract the local features of the image.

6 Appendix

6.1 Discretization calculation for scaling and rotation invariance of MEFMs

A digital image size of $N \times N$ is represented in the Cartesian coordinates, the computation of M_{nm}^{LPM} of MEFMs requires the conversion of the image between log-polar coordinates and Cartesian coordinates. Ping's CEM [44]

Table 6 Comparative results of the proposed approach and other published methods in recent years for D3 dataset corrupted by smoothing filter operation

CCPs (%)	KNN							
	N = 10 (lower-order moments)				N = 60 (higher-order moments)			
	EFMs	Fr-LMs	SOMs	Ours	EFMs	Fr-LMs	SOMs	Ours
Smoothing windows								
[3]	70.2	81.7	83.8	82.1	90.1	97.8	98.9	99.8
[5]	68.7	80.9	82.4	81.7	89.7	96.5	98.2	99.7
[7]	64.2	79.4	80.8	79.9	87.4	92.8	95.8	97.1
[9]	59.9	76.8	78.9	79.0	83.5	89.9	93.2	95.3

computing method will be used herein. We use $f_{LP}(\rho, \theta)$ to denote an image in log-polar coordinates, $f_c(x, y)$ to represent the image in Cartesian coordinates and $f[i, j]$ to denote the discrete coordinates. The log-polar coordinates (ρ, θ) can be converted to the Cartesian coordinates (x, y) using formula (33).

$$x = e^\rho \times \frac{N}{\sqrt{2}} \times \cos\theta; y = e^\rho \times \frac{N}{\sqrt{2}} \times \sin\theta \tag{33}$$

The Cartesian coordinates (x, y) can be changed into discrete coordinates $[i, j]$ with the following relationship:

$$i = -\lceil y \rceil + \frac{N}{2} + 1; j = \lceil x \rceil + \frac{N}{2} \tag{34}$$

Symbol $\lceil \cdot \rceil$ denotes the smallest integer not less than x or y . With the help of Eqs. (33) and (34), an image function $f_{LP}(\rho, \theta)$ in the log-polar coordinates can be achieved as

$$f_{LP}(\rho, \theta) = f \left[-\left\lceil e^\rho \times \frac{N}{\sqrt{2}} \times \sin\theta \right\rceil + \frac{N}{2} + 1, \left\lceil e^\rho \times \frac{N}{\sqrt{2}} \times \cos\theta \right\rceil + \frac{N}{2} \right] \tag{35}$$

We then let $\Delta\rho = \ln \frac{1}{K}$, $\Delta\theta = \frac{2\pi}{K}$, and $\rho_u = \ln \frac{u}{K}$, $u = 0, 1, \dots, K-1$; $\theta_v = \frac{2\pi v}{K}$, $v = 0, 1, \dots, K-1$.

Thus,

$$f_{LP}[\rho_u, \theta_u] \approx f \left[-\left\lceil e^{\rho_u} \times \frac{N}{\sqrt{2}} \sin\theta_v \right\rceil + \frac{N}{2} + 1, \left\lceil e^{\rho_u} \times \frac{N}{\sqrt{2}} \times \cos\theta_v \right\rceil + \frac{N}{2} \right] \tag{36}$$

In summary, integral Eq. (24) can be approximated as

$$\begin{aligned} M_{nm}^{LPM} &= \frac{1}{K^2} \sum_{u=0}^{K-1} \sum_{v=0}^{K-1} f_{LP}[\rho_u, \theta_v] w[u, v] e^{-j2n\pi\rho_u} e^{-jm\theta_v} \\ &= \frac{1}{K^2} \sum_{u=0}^{K-1} \sum_{v=0}^{K-1} G[u, v] e^{-j2n\pi \ln \frac{u}{K}} e^{-j2m\pi \frac{v}{K}} \end{aligned} \tag{37}$$

where, $G[u, v] = f_{LP}[u, v]w[u, v]$, and $w[u, v] = |T_n(\alpha, [(\frac{u-N/2}{N/2})^2 + (\frac{v-N/2}{N/2})^2]^{\frac{1}{2}}) | (\frac{u-N/2}{N/2})^2 + (\frac{v-N/2}{N/2})^2]^{\frac{1}{2}}$.

Eq. (37) shows the 2D discrete Fourier transform of $G[u, v]$. Therefore, the scaling and rotation invariant of

MEFMs for an image can be calculated by a 2D discrete Fourier transform (FFT).

Abbreviations

BFMs: Bessel–Fourier moments; CCPs: Correct classification percentages; EFMs: Exponent–Fourier moments; Fr-LMs: Fractional order Legendre moments; Fr-ZMs: Fractional order Zernike moments; GHMs: Gaussian–Hermite moments; LMs: Legendre moments; MEFMs: Modified exponent–Fourier moments; NaN: Not a number; OFMs: Orthogonal Fourier–Mellin moments; PCT: Polar cosine transform; PHT: Polar harmonic transform; PST: Polar sine transform; RST: Rotation, scaling, and translation; SNIRE: Statistical-normalization image-reconstruction error; SOM: Semi-orthogonal moments; ZMs: Zernike moments

Acknowledgements

The authors would like to thank the anonymous referees for their valuable comments and suggestions.

Authors’ contributions

JC developed the idea for the study modified exponent–Fourier moments and contributed the central idea in our manuscript; BH did the analyses for the properties of the proposed image moments, analyzed most of the data, and wrote the initial draft of the paper; and the remaining authors contributed to refining the ideas, carrying out additional analyses and finalizing this paper. All authors were involved in writing the manuscript. All authors read and approved the final manuscript.

Authors’ information

Bing He was born in 1982. He received the B.S. and M.S. degrees in Communication Engineering and Electrical Engineering from Northwestern Polytechnical University (NPU) and Shaanxi Normal University (SNNU), Xi’an, China in 2006 and 2009, respectively. Now he is an associate professor and pursuing the Ph.D. degree at Xidian University, Xi’an, China, and also a member of ACM. His research interests include image processing, object recognition, and digital watermarking.
 Jiang-Tao Cui was born in 1975. He received his B.S., M.S. and Ph.D. degrees in Computer Science and Technology from Xidian University, Xi’an, China in 1998, 2001, and 2005, respectively. He is currently a professor at the School of Computer Science and Technology in Xidian University, and also a member of IEEE. His research interests include image and video processing, pattern recognition, and high-dimensional indexing.
 Bin Xiao was born in 1982. He received his B.S. and M.S. degrees in Electrical Engineering from Shaanxi Normal University, Xi’an, China, in 2004 and 2007, and received his Ph.D. degree in computer science from Xidian University, Xi’an, China. He is now working as a professor at Chongqing University of Posts and Telecommunications, Chongqing, China. His research interests include image processing, pattern recognition, and digital watermarking.
 Yan-Guo Peng was born in 1986. He is currently a lecturer in Computer Architecture at Xidian University, Xi’an, China. He received his Ph. D. degree from Xidian University in 2016. His research interests include secure issues in data management, privacy protection, and cloud security.

Funding

This work was supported by National Natural Science Foundation of China (Grant No. 61472298, 61702403), the Fundamental Research Funds for the Central Universities (Grant No. JB170308, JBF180301), the Project funded by

China Postdoctoral Science Foundation (Grant No. 2018 M633473), Basic research project of Weinan science and Technology Bureau (Grant No. ZDYF-JCYJ-17), and by project of Shaan xi Provincial supports discipline (mathematics).

Availability of data and materials

- (1) The datasets generated in our experiments are available from Coil-20 and coil-100 image database and Butterfly images Database, URL link: <http://www.cs.columbia.edu/CAVE/databases/>. <http://cs.cqupt.edu.cn/info/1078/4189.htm>.
 (2) The datasets used or analyzed during the current study are available from the corresponding author on reasonable request.

Competing interests

The authors declare that they have no competing interests.

Author details

¹School of Computer Science and Technology, Xidian University, Xi'an 710071, China. ²School of Mathematics and Physics, Weinan Normal University, Weinan 714000, China. ³Chongqing Key Laboratory of Computational Intelligence, Chongqing University of Posts and Telecommunications, Chongqing 400065, China.

Received: 21 November 2018 Accepted: 13 June 2019

Published online: 18 July 2019

References

- B. Honarvar, R. Paramesran, C.L. Lim, Image reconstruction from a complete set of geometric and complex moments [J]. *Signal Processing* **98**(2), 224–232 (2014)
- B. Yang, M. Dai, Image reconstruction from continuous Gaussian–Hermite moments implemented by discrete algorithm [J]. *Pattern Recognition* **45**(4), 1602–1616 (2012)
- Y.D. Qu, C.S. Cui, S.B. Chen, et al., A fast subpixel edge detection method using Sobel–Zernike moments operator [J]. *Image & Vision Computing* **23**(1), 11–17 (2005)
- S. Ghosal, R. Mehrotra, A moment-based unified approach for image feature detection [J]. *IEEE Transactions on Image Processing*. A Publication of the IEEE Signal Processing Society **6**(6), 781–793 (1997)
- A. Hmimid, M. Sayyouri, H. Qjidaa, Fast computation of separable two-dimensional discrete invariant moments for image classification [J]. *Pattern Recognition* **48**(2), 509–521 (2015)
- Z. Shao, Y. Shang, Y. Zhang, et al., Robust watermarking using orthogonal Fourier–Mellin moments and chaotic map for double images [J]. *Signal Processing* **120**, 522–531 (2016)
- X.Y. Wang, Q.L. Shi, S.M. Wang, et al., A blind robust digital watermarking using invariant exponent moments [J]. *AEUE-International Journal of Electronics and Communications* **70**(4), 416–426 (2016)
- B. Xiao, G. Lu, Y. Zhang, et al., Lossless image compression based on integer discrete Tchebichef transform [J]. *Neurocomputing* **214**(C), 587–593 (2016)
- A. Khotanzad, Y.H. Hong, Invariant image recognition by Zernike moments [J]. *IEEE Transactions on Pattern Analysis & Machine Intelligence* **12**(5), 489–497 (2002)
- M.S. Choi, W.Y. Kim, A novel two stage template matching method for rotation and illumination invariance [J]. *Pattern Recognition* **35**(1), 119–129 (2002)
- H. Zhang, H. Shu, G. Coatrieux, et al., Affine Legendre moment invariants for image watermarking robust to geometric distortions [J]. *IEEE Transactions on Image Processing* **20**(8), 2189–2199 (2011)
- H. Zhang, H. Shu, G.N. Han, et al., Blurred image recognition by Legendre moment invariants [J]. *IEEE Transactions on Image Processing* **19**(3), 596–611 (2010)
- Deng AW, Wei CH, Gwo CY, Stable, fast computation of high-order Zernike moments using a recursive method [J]. *Pattern Recognition*, 2016, 56(C):16–25.
- C. Singh, R. Upneja, Accurate calculation of high order pseudo-Zernike moments and their numerical stability [J]. *Digital Signal Processing* **27**(1), 95–106 (2014)
- R.R. Galigekere, W. Holdsworth, M.N.S. Swamy, et al., Moment patterns in the Radon space [J]. *Optical Engineering* **39**(4), 1088–1097 (2000)
- B. Xiao, J.T. Cui, H.X. Qin, et al., Moments and moment invariants in the Radon space [J]. *Pattern Recognition* **48**(9), 2772–2784 (2015)
- C. Singh, A. Aggarwal, A comparative performance analysis of DCT-based and Zernike moments-based image up-sampling techniques [J]. *Optik-International Journal for Light and Electron Optics* **127**(4), 2158–2164 (2016)
- Z. Shao, H. Shu, J. Wu, et al., Quaternion Bessel–Fourier moments and their invariant descriptors for object reconstruction and recognition [J]. *Pattern Recognition* **47**(2), 603–611 (2014)
- J.S. Guf, W.S. Jiang, The Haar wavelets operational matrix of integration [J]. *International Journal of Systems Science* **27**(7), 623–628 (1996)
- H.J.A. Ferrer, I.D. Verduzco, E.V. Martinez, Fourier and Walsh digital filtering algorithms for distance protection [J]. *IEEE Transactions on Power Systems* **11**(1), 457–462 (1996)
- M. Hu, Visual pattern recognition by moment invariants [J]. *Information Theory Transactions on* **8**(2), 179–187 (1962)
- A. Abo-Zaid, O.R. Hinton, E. Horne, *About moment normalization and complex moment descriptors [C]// Paper presented at the International Conference on Pattern Recognition*. Berlin: Springer-Verlag Berlin Heidelberg, pp. 399–409 (1988)
- J. Flusser, On the independence of rotation moment invariants [J]. *Pattern Recognition* **33**(9), 1405–1410 (2000)
- Z. Ping, H. Ren, J. Zou, et al., Generic orthogonal moments: Jacobi–Fourier moments for invariant image description [J]. *Journal of Optoelectronics Laser* **40**(4), 1245–1254 (2007)
- Z. Ping, R. Wu, Y. Sheng, Image description with Chebyshev–Fourier moments [J]. *Journal of the Optical Society of America, Optics Image, Science, & Vision* **19**(9), 1748–1754 (2002)
- H. Ren, A. Liu, J. Zou, et al., *Character reconstruction with radial-harmonic-Fourier moments [C]// Paper presented at international conference on Fuzzy Systems and Knowledge Discovery* (IEEE Computer Society, 2007), pp. 307–310
- G. Gao, G. Jiang, Bessel–Fourier moment-based robust image zero-watermarking [J]. *Multimedia Tools & Applications* **74**(3), 841–858 (2015)
- B. Xiao, G.Y. Wang, W.S. Li, Radial shifted Legendre moments for image analysis and invariant image recognition [J]. *Image & Vision Computing* **32**(12), 994–1006 (2014)
- B. Fu, J. Zhou, Y. Li, et al., Image analysis by modified Legendre moments [J]. *Pattern Recognition* **40**(2), 691–704 (2007)
- R. Mukundan, S.H. Ong, P.A. Lee, Image analysis by Tchebichef moments [J]. *IEEE Transactions on Image Processing* **10**(9), 1357–1364 (2001)
- P.T. Yap, R. Paramesran, S.H. Ong, Image analysis by Krawtchouk moments [J]. *IEEE Transactions on Image Processing* **12**(11), 1367–1377 (2003)
- P.T. Yap, R. Paramesran, S.H. Ong, Image analysis using Hahn moments [J]. *IEEE Transactions on Pattern Analysis & Machine Intelligence* **29**(11), 2057–2062 (2007)
- H. Zhu, H. Shu, J. Liang, et al., Image analysis by discrete orthogonal Racah moments [J]. *Signal Processing* **87**(4), 687–708 (2007)
- B. Xiao, L. Li, Y. Li, et al., Image analysis by fractional-order orthogonal moments [J]. *Information Sciences* **382–383**, 135–149 (2017)
- J.W. Yang, D.J. Jin, Lu Z D, Fractional order Zernike moment [J]. *Journal of Computer-Aided Design & Computer Graphics* **29**(3), 480–484 (2017)
- B. Chen, M. Yu, Y.Q. Su, et al., Fractional quaternion Zernike moments for robust color image copy-move forgery detection [J]. *IEEE Access* **6**(C), 56637–56646 (2018)
- B.J. Chen, X.M. Qi, X.M. Sun, et al., Quaternion pseudo-Zernike moments combining both of RGB information and depth information for color image splicing detection [J]. *Journal of Visual Communication and Image Representation* **49**(C), 283–290 (2017)
- X. Wang, T. Yang, F. Guo, Image analysis by circularly semi-orthogonal moments [J]. *Pattern Recognition* **49**(C), 226–236 (2016)
- H.T. Hu, Y.D. Zhang, et al., Orthogonal moments based on exponent functions: exponent-Fourier moments [J]. *Pattern Recognition* **47**(8), 2596–2606 (2014)
- B. Xiao, J.F. Ma, X. Wang, Image analysis by Bessel–Fourier moments [J]. *Pattern Recognition* **43**(8), 2620–2629 (2010)
- P.T. Yap, X. Jiang, A.C. Kot, Two-dimensional polar harmonic transforms for invariant image representation [J]. *IEEE Transactions on Pattern Analysis & Machine Intelligence* **32**(7), 1259–1270 (2010)
- Coil-20 and coil-100 image database, <http://www.cs.columbia.edu/CAVE/databases>. Accessed 30 June 2017
- B. Xiao, G. Lu, T. Zhao, et al., Rotation, scaling, and translation invariant texture recognition by Bessel–Fourier moments [J]. *Pattern Recognition & Image Analysis* **26**(2), 302–308 (2016)
- Z.L. Ping, Y.J. Jiang, S.H. Zhou, et al., *FFT algorithm of complex exponent moments and its application in image recognition* (Proceedings of the SPIE International Society for Optical Engineering, Athens, 2014)
- N. Ohtsu, A threshold selection method from gray-level histograms [J]. *IEEE Transactions on Systems Man & Cybernetics* **9**(1), 62–66 (1979)

46. Butterfly images Database, <http://cs.cqupt.edu.cn/info/1078/4189.htm>. Accessed 15 May 2015
47. J.W. Han, X. Ji, X.T. Hu, et al., Representing and retrieving video shots in human-centric brain imaging space [J]. *IEEE Transactions on Image Processing* **22**(7), 2723–2736 (2013)
48. T. Zhang, L. Guo, K.M. Li, et al., Predicting functional cortical ROIs via DTI-derived fiber shape models [J]. *Cerebral cortex* **22**(4), 854–864 (2012)

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- ▶ Convenient online submission
- ▶ Rigorous peer review
- ▶ Open access: articles freely available online
- ▶ High visibility within the field
- ▶ Retaining the copyright to your article

Submit your next manuscript at ▶ [springeropen.com](https://www.springeropen.com)
