Pattern Recognition 48 (2015) 3895–3903

Contents lists available at ScienceDirect

Pattern Recognition

journal homepage: www.elsevier.com/locate/pr

Quaternion generic Fourier descriptor for color object recognition

Heng Li *, Zhiwen Liu, Yali Huang, Yonggang Shi

School of Information and Electronics, Beijing Institute of Technology, Beijing 100081, China

ARTICLE INFO

Article history:
Received 6 November 2014
Received in revised form
24 April 2015
Accepted 1 June 2015
Available online 17 June 2015

Keywords:
Quaternion
Generic Fourier descriptor
Invariant descriptor
Color object recognition

ABSTRACT

By introducing quaternions, Fourier transforms, the fundamental tool in signal and image processing, were extended to treat a color image in a holistic manner. On the other hand, color image description has long been a major topic of pattern recognition. In this work, inspired by the quaternion Fourier transforms, we propose a quaternion generic Fourier descriptor to describe color images holistically and effectively without losing any color information. The theoretical framework to construct a set of invariants with respect to geometric transformation is also provided. Moreover, several experiments are conducted on color image databases to illustrate the effectiveness of the proposed invariants. The results provide supports for that the quaternion generic Fourier descriptor invariants have striking robustness to geometric transformation and noise. Remarkable progress in accuracy, robustness and efficiency of color object recognition has been made by the proposed descriptor. In addition, it presents superior performances on real recognition scenarios.

1. Introduction

Since the discovery of the fast Fourier transform in 1965, Fourier transforms have been a fundamental tool in signal and image processing, and widely used in pattern recognition and computer vision [1–3]. Within the last few decades, new imaging techniques have enabled the acquisition of color images with unprecedented accuracy and efficiency. There is an urgent demand for effective tools to facilitate the processing of color images. Nevertheless, the conventional image processing methods are limited by Fourier transform, which cannot tackle color images directly. Two frameworks are often applied in the pioneering work of color image processing. The first one transforms the color image into gray-scale [4], which consequently loses some color information, and the other decomposes the color image into three channels, then processes these three channels separately [5–7].

To overcome the limitation, the idea of computing the Fourier transform of a color image has recently been realized [8]. Hypercomplex numbers, in particular, quaternions have been employed in color image processing. Quaternions were introduced by Hamilton in 1843, and in recent years the first definition of a quaternion Fourier transform (QFT) was discovered by Ell in his Ph.D. dissertation. Then Sangwine first derived quaternion discrete Fourier transform (QDFT) and utilized quaternions to process digital color images [8–10]. As it handles the color image as a vector field holistically rather than separate components, quaternion Fourier transforms have recently enlightened the color image processing, and have gained popularity in image denoising [11,12], color image registration [13,14], watermarking [15] and saliency detection [16–18].

Image description aims to describe or represent the image by certain features that can be used to recognize a similar image or object from specific databases. Such studies have long been a major topic of pattern recognition because of the difficulties in developing the discriminative invariants with respect to geometric transformation, namely rotation, scaling and translation (RST). Hu [19] published the first significant paper on the use of image moment invariants, which is invariant under geometric transformations. Then Teague [20] proposed Legendre moments and Zernike moments by replacing transform kernel of geometric moment with Legendre polynomial and Zernike polynomial respectively. Based on the polar transformation and the Fourier–Mellin transform, the Fourier–Mellin descriptor was introduced to process image by Sheng and Shen [21]. Nevertheless, the invariants mentioned above focus on gray-scale images. When it comes to color images more problems arise, such as, how to preserve the color information and how to process the color image holistically. By introducing algebra of quaternion, several types of quaternion-based invariants have recently been developed to tackle color images. In 2011, Guo and Zhu [22] published their work on the quaternion Fourier–Mellin moment invariants and tested their invariance with respect to rotation, scaling and translation transformation. A hypercomplex polar Fourier descriptor, which can be utilized to represent and reconstruct images, was proposed by Yang and Kamata [23] using basis functions. And in the following research, they reported the fast computational version of the descriptor [24]. Chen et al. [25] extended the conventional Zernike

* Corresponding author.
E-mail address: heng.li@bit.edu.cn (H. Li).

http://dx.doi.org/10.1016/j.patcog.2015.06.002
0031-3203/

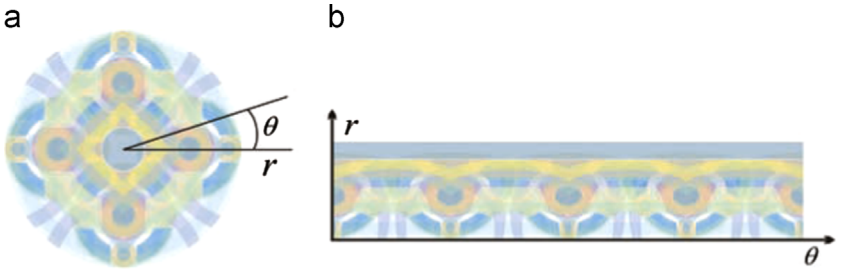

Fig. 1. (a) Original image in polar space and (b) the polar image of (a) plotted into Cartesian space.

moment defined in gray-scale image to color image using the algebra of quaternions in 2012. Recently, Shao et al. [26] presented the quaternion Bessel–Fourier moments which have a good performance on object reconstruction.

As an alternative to the moment-based descriptors, the generic Fourier descriptor (GFD) was proposed by Zhang and Lu in 2002 [27] to overcome some the shortcomings of Zernike moment descriptor, such as the complex computation, the inconsistency of the radial features and circular features. Since then, motivated by its efficient computation and robustness, GFD has been applied to a number of pattern recognition and computer vision problems [28–30]. In this paper, the generic Fourier descriptor is extended from complex numbers to quaternion algebra, and applied to color images.

The first two contributions of this paper are the construction of a quaternion generic Fourier descriptor (QGFD), and the development of robust QGFD invariants with respect to geometric transformation. The experimental results illuminate the outstanding robustness of QGFD invariants to geometric transformation and noise. The third contribution is the advance made by the proposed descriptor in color object recognition. QGFD outperforms the existing color image descriptors in the aspects of accuracy, robustness and efficiency of color object recognition. On the applications of color image retrieval and face recognition, it also provides state-of-the-art performances.

The remainder of this paper is organized as follows. In Section 2, we introduce the quaternion generic Fourier descriptor and its invariants for ration, translation and scaling of color image. Several experiments and discussions are provided in Section 3 to illustrate the superior performance of the proposed descriptors. Section 4 concludes this paper.

2. Quaternion generic Fourier descriptor

2.1. Quaternion number

The quaternions are a number system that extends the complex numbers. They were first introduced by the mathematician Hamilton in 1843. Then Sangwine [8–10] applied them in digital color image processing. By encoding the three channel components on the imaginary parts, quaternions treat the color image holistically as a vector field. A quaternion q comprises one real part and three imaginary parts as follows:

$$q = a + ib + jc + kd, \tag{1}$$

where a, b, c, d are real numbers, and i, j, k are orthogonal imaginary units obeying the following rules:

$$i^2 = j^2 = k^2 = -1, \tag{2}$$

$$ij = -ji = k, \quad jk = -kj = i, \quad ki = -ik = j. \tag{3}$$

From Eq. (3), it can be found that the quaternion multiplication is not commutative.

The conjugate of a quaternion is specified as

$$\overline{q} = a - in - jc - kd, \tag{4}$$

and for any two quaternions p and q, $\overline{p \cdot q} = \overline{q} \cdot \overline{p}$.

The modulus of a quaternion is given by

$$|q| = \sqrt{a^2 + b^2 + c^2 + d^2}. \tag{5}$$

2.2. Quaternion generic Fourier transform and descriptor

Let $f(r,\theta)$ be an RGB image defined in polar coordinates, then a quaternion-based model for color image can be represented as

$$f(r,\theta) = if_R(r,\theta) + jf_G(r,\theta) + kf_B(r,\theta), \tag{6}$$

where $f_R(r,\theta)$, $f_G(r,\theta)$, and $f_B(r,\theta)$ represent respectively the red, green and blue components of the pixel.

Inspired by the original generic Fourier transform [27], the first step is to transform the polar image in polar space to a ordinary two-dimensional rectangular image in Cartesian space (shown in Fig. 1). Then 2-D quaternion discrete Fourier transform is applied on this rectangular image. Because of the non-commutative property of quaternion multiplication, the left-side and right-side alterations of quaternion polar Fourier transform are obtained as

$$QF^L_{R,T}(\rho,\phi) = \sum_r \sum_n e^{\mu 2\pi((r/R)\rho + (2\pi n/T)\phi)} f(r,\theta_n), \tag{7}$$

$$QF^R_{R,T}(\rho,\phi) = \sum_r \sum_n f(r,\theta_n) e^{\mu 2\pi((r/R)\rho + (2\pi n/T)\phi)}, \tag{8}$$

where R and T determine the resolution of radial frequency and angular frequency respectively $0 \le r < R$ and $\theta_n = n \times (2\pi/T)$ $(0 \le n < T)$, $0 \le \rho < R$, $0 \le \phi < T \cdot \mu$ is a unit pure quaternion defined as $\mu = (i+j+k)/\sqrt{3}$. here.

As $\overline{p \cdot q} = \overline{q} \cdot \overline{p}$, the relationship between the left-side and right-side alterations for the same color image is

$$\begin{aligned} QF^L_{R,T}(\rho,\phi) &= \sum_r \sum_n e^{\mu 2\pi((r/R)\rho + (2\pi n/T)\phi)} f(r,\theta_n) \\ &= \overline{\sum_r \sum_n f(r,\theta_n) e^{-\mu 2\pi((r/R)\rho + (2\pi n/T)\phi)}} \\ &= -\overline{\sum_r \sum_n f(r,\theta_n) e^{-\mu 2\pi((r/R)\rho + (2\pi n/T)\phi)}} \\ &= -\overline{QF^R_{R,T}(-\rho,-\phi)}. \end{aligned} \tag{9}$$

Eq. (9) indicates that the left-side and right-side alterations are equivalent to describe color images. In the following part, the alteration of quaternion polar Fourier transform refers to left-side type. The quaternion generic Fourier descriptor is defined as

$$\begin{aligned} QD = \{ &\|QF(0,0)\|, \|QF(0,1)\|, \ldots, \|QF(0,T)\|, \\ &\ldots, \|QF(R,0)\|, \ldots, \|QF(R,T)\| \}. \end{aligned} \tag{10}$$

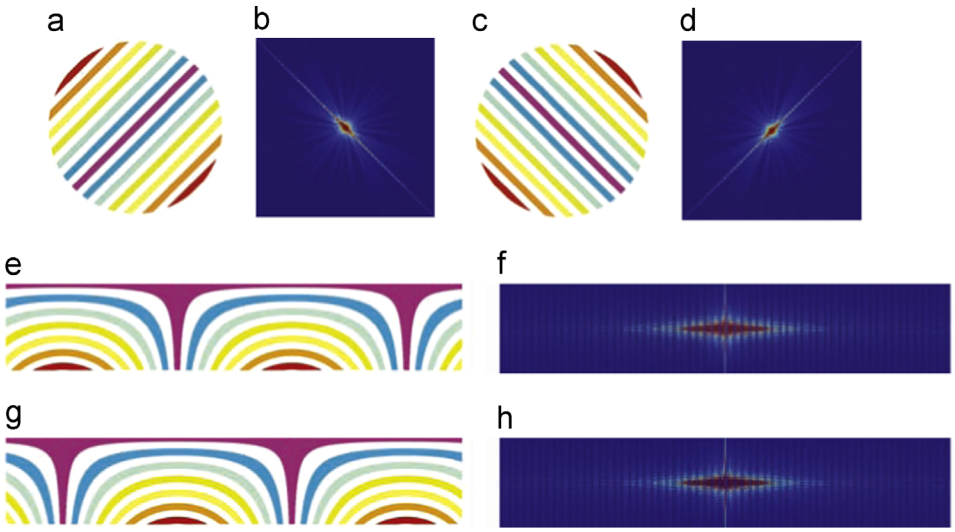

Fig. 2. (a) An example image, (b) Fourier spectra distributions of (a), (c) image (a) rotated by 90°, (d) Fourier spectra distributions of (c), (e) the polar image of (a), (f) Fourier spectra distributions of (e), (g) the polar image of (c), and (h) Fourier spectra distributions of (g).

2.3. The invariants of quaternion generic Fourier descriptor

In the following part, we provide the framework to construct the geometric transformation invariants.

2.3.1. Rotation invariance

As well as 2-D Fourier transform, the features captured from 2-D QFT are not invariant to rotation (Fig. 2(a)–(d)). As demonstrated in Fig. 2, due to polar transformation, the rotation of original image is transformed into translation in the polar image. Let $f(r,\theta+\varphi)$ denote the rotated version of a color image $f(r,\theta)$ by the angle φ, the quaternion generic Fourier descriptors of $f(r,\theta+\varphi)$ and $f(r,\theta)$ have the following relations:

$$\begin{aligned} QF'_{R,T} &= \sum_r\sum_n e^{\mu 2\pi((r/R)\rho+(2\pi n/T)\phi)} f(r,\theta_n+\varphi) \\ &= \sum_r\sum_n e^{\mu 2\pi(-\varphi\phi)} e^{\mu 2\pi((r/R)\rho+((2\pi n/T)+\varphi)\phi)} f(r,\theta_n+\varphi) \\ &= e^{-\mu 2\pi\varphi\phi}\sum_r\sum_n e^{\mu 2\pi((r/R)\rho+\theta'_n\phi)} f(r,\theta'_n) \\ &= e^{-\mu 2\pi\varphi\phi} QF_{R,T}, \end{aligned} \tag{11}$$

where $Q'_{R,T}$ and $Q_{R,T}$ are the results of $f(r,\theta+\varphi)$ and $f(r,\theta)$ that underwent quaternion generic Fourier transform.

Taking the norm on both sides of Eq. (11)

$$\begin{aligned} \|QF'_{R,T}\| &= \left\|e^{-\mu 2\pi\varphi\phi} QF_{R,T}\right\| \\ &= \left\|e^{-\mu 2\pi\varphi\phi}\right\| \cdot \|QF_{R,T}\| \\ &= \|QF_{R,T}\|. \end{aligned} \tag{12}$$

Further, we have the QGFD of $f(r,\theta+\varphi)$ equal to that of $f(r,\theta)$.

2.3.2. Translation invariance

Translation invariance is achieved by putting the origin of coordinates at the common centroid. Suk and Flusser have depicted the common centroid (x_c,y_c) of all three channels as follows [6]:

$$x_c = \frac{(m_{1,0}(f_R)+m_{1,0}(f_G)+m_{1,0}(f_B))}{m_{0,0}}, \tag{13}$$

$$y_c = \frac{(m_{0,1}(f_R)+m_{0,1}(f_G)+m_{0,1}(f_B))}{m_{0,0}}, \tag{14}$$

$$m_{0,0} = m_{0,0}(f_R)+m_{0,0}(f_G)+m_{0,0}(f_B), \tag{15}$$

where $m_{0,0}(f_R)$, $m_{1,0}(f_R)$ and $m_{0,1}(f_R)$ are respectively the zero-order geometric moment and first-order ones for R channel, $m_{0,0}(f_G)$, $m_{1,0}(f_G)$ and $m_{0,1}(f_G)$ for G channel, and $m_{0,0}(f_B)$, $m_{1,0}(f_B)$ and $m_{0,1}(f_B)$ for B channel.

2.3.3. Scaling invariance

Scale invariance is produced by altering each image, so that its zero-order geometric moment is normalized. Namely, $m_{0,0}=\beta$, where β is a predetermined value.

Through the framework mentioned above, the invariants of QGFD (QGFDIs) with respect to geometric transform are constructed.

3. Experiments and results

In this section, several experiments are provided to validate the effectiveness of the proposed QGFD invariants under geometric transformation, and their robustness to different kinds of noises. The performance comparisons between the proposed descriptor and the existing methods for color object recognition are also presented. Furthermore, the proposed descriptor is applied to color image retrieval of real scene images and highly textured images as well as face recognition.

3.1. Experiments on geometric transformation invariance and noise robustness

3.1.1. Geometric transformation invariance

With the purpose of evaluating the invariance property of the proposed QGFDIs to geometric transformation, an image with size 128×128 arbitrarily selected from the Columbia University Image Library (COIL-100) database [31] was utilized. The size of the original image was regulated to 200×200 by adding some background pixels, in order to obtain the entire transformed image after transformation.

For this experiment, the original image (Fig. 3(a)) has undergone different geometric transformations (Fig. 3(b)–(h)). The angular resolution was set to 360 (sampling at every degree) and the radial resolution depended on the image size (about half length of the image). Therefore, the proposed QGFDIs with $R=100$, $T=360$, was calculated for images. Table 1 shows the value of the first nine QGFD invariants and that of σ/μ, where μ denotes the mean of invariants and σ is the standard deviation.

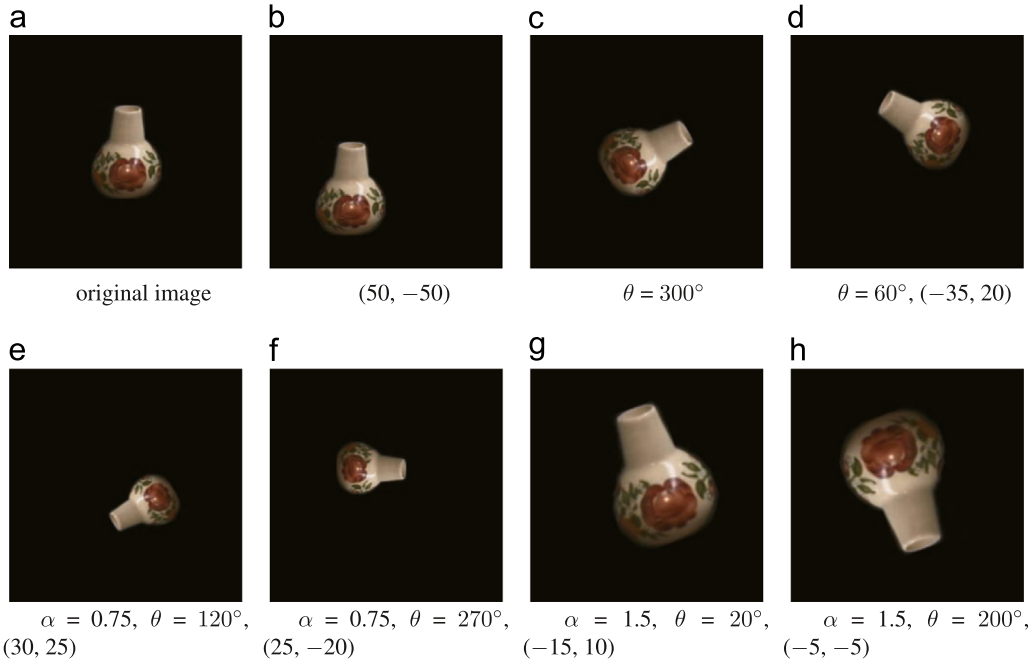

Fig. 3. Geometric transformed images of image (a), α is the scaling factor, θ is the rotation angle, (x, y) represents the translation along the x-axis and the y-axis.

Table 1
Value of QGFDIs of images shown in Fig. 3.

Indexes	QD_1	QD_2	QD_3	QD_4	QD_5	QD_6	QD_7	QD_8	QD_9
(a)	75.140	19.923	9.5985	5.7608	2.9448	3.1765	2.9503	2.3126	2.0873
(b)	74.981	19.727	9.5104	5.7307	3.0227	3.1474	2.9410	2.2457	2.0527
(c)	75.045	19.769	9.5104	5.8293	2.9714	3.1334	3.0076	2.3265	2.0612
(d)	74.956	19.698	9.5072	5.7360	3.0259	3.1421	2.9559	2.2395	2.0510
(e)	75.182	19.998	9.]5780	5.7282	2.9437	3.2032	2.8743	2.3079	2.0924
(f)	74.943	19.640	9.5792	5.8082	3.0014	3.1277	3.0510	2.2698	2.0597
(g)	75.068	19.768	9.6333	5.8559	3.0071	3.1372	2.9474	2.3399	2.0781
(h)	75.034	19.790	9.5419	5.7157	2.9525	3.1906	2.9893	2.2404	2.0743
σ/μ	0.0011	0.0059	0.0048	0.0092	0.0116	0.0091	0.0177	0.0180	0.0076

It can be seen from Table 1 that σ/μ of QGFDIs is less than 0.0180. Besides, we calculated the Euclidean distances between the original image and the transformed images, as well as the Euclidean distances between the original image and other images of COIL-100. The p-value of these two Euclidean distance sets is 3.12×10^{-11}. Consequently, QGFDIs remain stable under different geometric transformations, and can be a useful tool to describe color images.

3.1.2. Noise robustness

Another image with size 128×128 has been randomly selected from COIL-100 [31] to test the robustness of QGFD. The actual size of the original image was changed to 150×150.

To test the robustness of the proposed QGFDIs, the original image (Fig. 4(a)) underwent adding white Gaussian noise with different standard deviations and salt-and-pepper noise with different noise densities respectively (Fig. 4(b)–(h)). The value of QGFDIs with $R=70$, $T=360$ is provided in Table 2.

Table 2 illuminates that excellent robustness to various noise has been achieved. σ/μ of QGFDIs in the cases of noise-free and different types of noise is less than 0.0775. The Euclidean distance sets from the original image to the images with noise and to the other images of COIL-100 were calculated, and the p-value of them is 6.35×10^{-6}. Therefore, QGFDIs are robust and efficient as color image descriptor invariants.

3.2. Color object recognition

With intent to compare the QGFDIs with the existing quaternion invariants, we carry out experiments on color object recognition with two popular moment-based quaternion invariants, including quaternion Fourier–Mellin moment invariants (QFMMIs) [22] and quaternion Zernike moment invariants (QZMIs) [25].

3.2.1. Performance comparison in COIL-100

This experiment is conducted under the completely same conditions to Ref. [25]. The training set were the 14 images with size 128×128 selected from COIL-100 [31] following Ref. [25] (Fig. 5). After adding some background pixels, the actual size of the original image was 204×204. Then the testing set comprised 1176 ($=14 \times 12 \times 7$) images, which were obtained by translating, rotating and scaling each image with $x=-11$, $y=9$, $\theta \in \{0, 30, \ldots, 300, 330\}$ and $\alpha \in \{0.5, 0.75, \ldots, 1.75, 2.0\}$ respectively. White Gaussian noise and salt-and-pepper noise were added to the testing images as well. Subsequently, QGFDIs were diverted with $R=50$, $T=60$.

For $R=50$, $T=60$, the total number of QGFDIs is 3000. This is a comparatively large feature vector. Thus we adopted dimension reduction techniques to limit the size of feature vector. Principal component analysis (PCA) was firstly applied. As a result of PCA, the first 12 principal components satisfied 90% of the total variance explained; the first 6 principal components satisfied 80%; the first

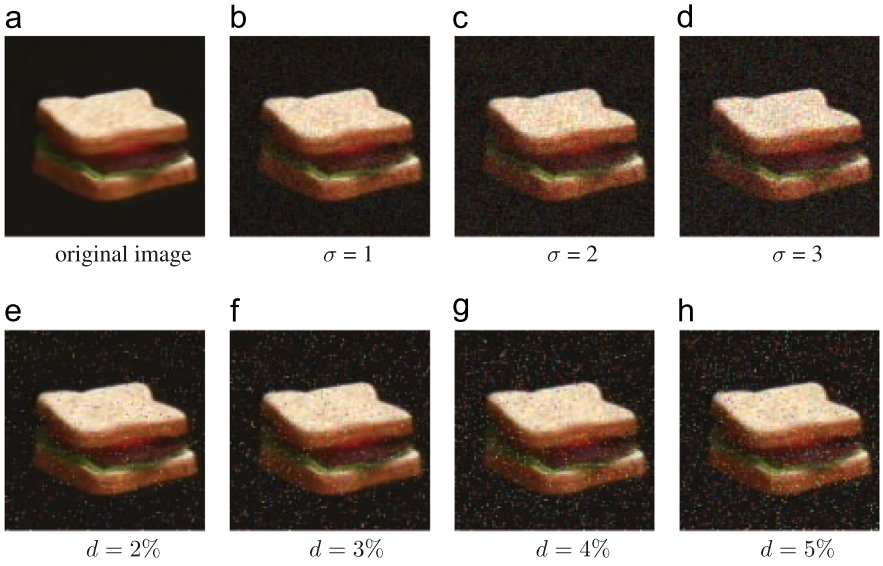

Fig. 4. Examples of image with varying noise. σ is the STD of Gaussian noise, and d is the density of salt-and-paper noise.

Table 2
Value of QGFDIs of images shown in Fig. 4.

Indexes	QD_1	QD_2	QD_3	QD_4	QD_5	QD_6	QD_7	QD_8	QD_9
(a)	72.978	19.594	7.7665	6.3721	4.3276	3.3781	3.0686	2.6953	2.0450
(b)	72.598	19.432	7.5895	6.3218	4.1449	3.2788	2.9923	2.5992	1.9509
(c)	72.474	19.127	7.7218	6.1994	4.2578	3.3550	2.9091	2.7584	1.9957
(d)	72.385	19.305	7.5355	6.3206	4.1173	3.2646	3.0465	2.4402	1.8784
(e)	72.847	19.524	7.8588	6.5213	4.3509	3.4030	3.2000	2.6784	2.1681
(f)	72.322	19.126	7.5632	6.2279	4.2661	3.3042	2.9960	2.6044	2.0600
(g)	72.597	19.311	7.6849	6.2522	4.2672	3.2939	3.1416	2.5750	2.2815
(h)	72.012	19.059	7.2093	6.1613	3.8033	3.1418	2.8923	2.1286	2.0001
σ/μ	0.0042	0.0102	0.0259	0.0182	0.0421	0.0247	0.0351	0.0775	0.0619

Fig. 5. Some examples of color objects selected from the Columbia University Image Library database.

Table 3
Recognition rates (%) of QFMMIs, QZMIs and QGFDIs for object recognition with geometric transformation and noise.

Methods	Noise-free	Gaussian noise			Salt-and-pepper noise		
		σ=5	σ=7	σ=9	d=1%	d=2%	d=3%
QFMMIs	100.00	100.00	96.60	89.63	98.64	88.95	75.85
QZMIs	100.00	100.00	100.00	96.26	99.91	92.77	84.35
QGFDIs	100.00	100.00	100.00	100.00	100.00	98.64	98.04
PCA-70%	100.00	100.00	100.00	99.66	100.00	98.04	95.07
PCA-80%	100.00	100.00	100.00	99.68	100.00	98.30	95.15
PCA-90%	100.00	100.00	100.00	100.00	100.00	98.55	97.96
QGFDIs-50	100.00	100.00	99.87	99.32	96.94	89.88	85.12
QGFDIs-60	100.00	100.00	100.00	100.00	100.00	93.62	90.22
QGFDIs-70	100.00	100.00	100.00	100.00	100.00	94.13	92.86

4 principal components satisfied 70%. Accordingly, only a few invariants were significant. Alternatively, the energy compaction property of Fourier transform has been widely applied to signal compression. Likewise, as QGFDIs were arranged in the order from low to high frequency, we briefly used the first 50, 60, 70 invariants of QGFDIs as feature vectors to limit the dimension. The original invariants of QGFDIs as well as the dimension reduced invariants were finally classified by the minimum-Euclidean-distance.

Table 3 summarizes the correct classification percentage (CCP) of QGFDIs with various numbers of invariants. And the results of QFMMIs and QZMIs reported in Ref. [25] are provided as well. The number of invariants used in Ref. [25] of the methods are as follows: The 24 QFMMIs are the invariants defined in Eq. (15) of Ref. [22] with order from 0 to 4. The 20 QZMIs are the invariants depicted in Eq. (34) of Ref. [25] with order from 1 to 4, and calculated with the modified Kintner's method [32]. Then they are classified by the minimum-Euclidean-distance as well.

QGFDIs overcome the existing moment invariants in the strong noise cases. With the increase of noise, the robustness of QGFDIs enables reasonable recognition results. PCA performs outstandingly that it allows us to reduce the dimension of feature vector strikingly

Fig. 6. Some examples of color objects selected from the Amsterdam Library of Object Images database.

Table 4
Recognition rates (%) of QFMMIs, QZMIs and QGFDIs for object recognition with geometric transformation and noise.

Types and levels of noise	QFMMIs	QZMIs	QGFDIs
Noise-free	99.89	99.86	100.00
Gaussian noise with STD = 0.1	83.11	88.70	99.51
Gaussian noise with STD = 0.2	75.67	84.92	99.40
Gaussian noise with STD = 0.3	68.56	81.99	98.93
Gaussian noise with STD = 0.4	61.83	79.49	98.79
Gaussian noise with STD = 0.5	57.45	79.02	98.29
Salt-and-pepper noise with noise density = 0.1%	83.59	86.32	99.49
Salt-and-pepper noise with noise density = 0.2%	75.46	80.01	98.99
Salt-and-pepper noise with noise density = 0.3%	68.44	74.07	98.66
Salt-and-pepper noise with noise density = 0.4%	61.26	69.57	98.36
Salt-and-pepper noise with noise density = 0.5%	54.91	66.58	98.14
Elapsed time (s)	1.518	2.732	0.801

without sacrificing the recognition performance too much. On the other hand, the invariants of low frequency also provide discriminative features of objects. The performances of the invariants of low frequency are comparable with those of all invariants, particularly in the low noise cases. Hence for the sake of simplification, the dimension of QGFDIs can be briefly reduced by discarding the invariants of high frequency.

3.2.2. Performance comparison in large data volume

To further verify the performance of the proposed QGFDIs in a larger data volume, a set of 1000 color images with size of 192×144 was selected from the public Amsterdam Library of Object Images (ALOI) database [33] and utilized in this experiment (shown in Fig. 6). QFMMIs and QZMIs were carried out with the number of invariants as presented in Section 3.2.1. QGFDIs were performed with $R=50$, $T=60$, and only the first 20 invariants of each image were used as the feature vector to limit the dimension.

To contain the entire transformed image, the tested images have been added background to 400×400. The 1000 original images was treated as the training set. Then the testing set was constructed as follows: the original images were scaled with the factor $\alpha \in \{0.75, 1.0, 1.5\}$ and rotated from 0° to 300°, with an increment 60°. Each image was then translated with (x,y), whose two elements were randomly selected from an interval, $[-60, 60]$. Therefore, the testing set was composed of 18 000 ($=1000 \times 3 \times 6$) images. This was followed by adding white Gaussian noise and salt-and-pepper noise.

The CCP was measured by the Euclidean distance based nearest neighbor classifier. Besides, the average elapsed time in Matlab 2012b on a 3.1 GHz Core PC station (Lenovo ThinkCentre M8400t) of computing 18 images (every selected image was transformed to 18 test images as described above) was also recorded.

The recognition results of QFMMIs, QZMIs and the proposed QGFDIs for both noise-free and noisy cases are shown in Table 4. The results demonstrate that (i) all invariants perform quite good (CCP $\geq 99\%$) in noise-free case. With the noise level going up, all of the recognition rates decrease. (ii) In contrast with the experiment in Section 3.2.1, the vast volume of image database and the numerous classes make the recognition of color objects a burdensome task. Thus the misrecognition occurs even in noise-free cases of QFMMIs and QZMIs. Moreover, the CCP of QFMMIs and QZMIs decreases severely once noise is added into the experiments. (iii) Due to the more discriminative power and the superior robustness, QGFDIs not only correctly recognize all image in the noise-free case, but also present a competitive performance whatever the noise type and the noise level is. (iv) As a consequence of its straightforward computation, the average elapsed time of QGFDIs is quite shorter than that of QFMMIs and QZMIs.

The computation of QGFDIs is much easier than the moment-based descriptors. Assume the size of an RGB image to be $N \times N$ pixels and let n be the maximum order of QFMMIs and QZMIs. The computation of QFMMIs needs $O\left(n \times N^2\right)$ multiplications and $O\left(n \times N^2\right)$ additions. Using the modified Kintner's method [25], there is no more factorial function involved computation of QZMIs. For the same image, the computation of QZMIs still needs $O\left(n^2 \times N^2\right)$ multiplications and $O\left(n^2 \times N^2\right)$ additions. Since all QGFDIs are obtained by the identical transform with specific R and T, the computation of QGFDIs is $O(R \times T)$ multiplications and $O(R \times T)$ additions. Owing to the relatively straightforward computation, QGFDIs circumvent the time-consuming computation of various order moment invariants.

3.3. Experiments on real recognition scenarios

We next apply the proposed descriptor to real recognition scenarios, such as color image retrieval of real scene images and face recognition.

3.3.1. Color image retrieval

In this experiment, we verify the retrieval performance of our descriptor for real scene images. As illuminated in Fig. 7, 50 images were separately selected from PASCAL VOC 2007 database [34] and Texture database [35]. In addition, various radial and angular resolutions were applied to demonstrate the relationship between parameters and performances.

Each image of the databases was transformed using following stochastic parameters, scale factor $\alpha \in \{0.8, 0.9, 1.0, 1.1, 1.2\}$, rotation factor $\theta \in [0, 360)$ and translation factors $x, y \in [-20, 20]$. Subsequently, the transformed image was randomly added by white Gaussian noise with $\sigma \in \{0, 1, 2, 3\}$ or salt-and-pepper noise with $d \in \{0, 0.01, 0.02, 0.03\}$. Each image was thus multiplied to 60 images by repeating such process, and one of them was arbitrarily selected as training data. Therefore, both databases had 50 images for training and 2950 ($=50 \times 59$) images for testing. QGFDIs with various radial and angular resolutions were computed. The minimum-Euclidean-distance was used as the classifier for the invariants.

Tables 5 and 6 respectively present the recognition results of PASCAL VOC 2007 and Texture databases. For the both databases, QGFDIs perform promisingly with high enough radial and angular

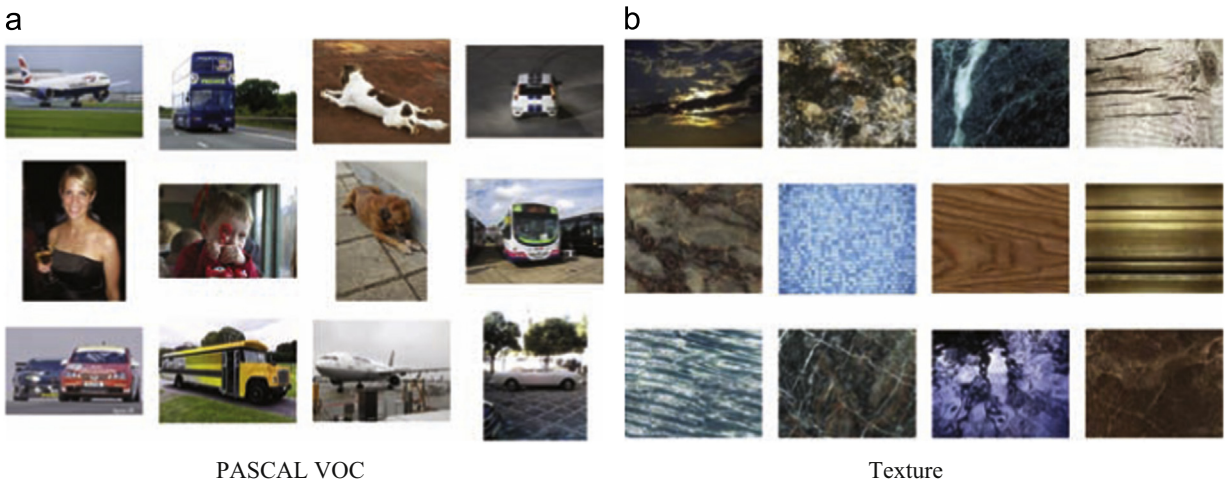

Fig. 7. Some examples of color images selected from PASCAL VOC 2007 and Texture database.

resolutions that QGFDIs with $R=50$, $T=60$ are sufficient to describe most images. The diversity of real scene image size impacts the performance of QGFDIs, especially for low radial and angular resolutions. And it demonstrates that QGFDIs are competent to describe highly textured images as well. Besides the low angular resolution limits the performance of QGFDIs significantly, while they are comparatively robust to the low radial resolution.

3.3.2. Face recognition

For the purpose of substantiating the performance of our descriptor for face recognition, an experiment is conducted on three color face databases provided by University of Essex [36], including Faces 95, Faces 96 and Grimace as shown in Fig. 8. Faces 95 comprises 72 individuals, and 20 images per individual with size of 180×200 and variation of backgrounds, head scale, position and lighting. Faces 96 comprises 152 individuals, and 20 images per individual with size of 196×196 and variation of complex backgrounds, head scale, position and lighting. Grimace comprises 18 individuals, and 20 images per individual with size of 180×200 and variation of head pose and expression.

The feature vector was the first 20 invariants of QGFDIs with $R=50$, $T=60$. Then support vector machine (SVM) with the kernel of Gaussian radial basis function was applied to recognize the faces. For each individual, 16 images were randomly selected as the training data and the else were the test data. And the recognition experiments were repeated for 10 times independently. Table 7 shows the face recognition results of mean recognition precision (TrPr, TePr), recall rate (TrRe, TeRe) and *F*-score (TrF, TeF) of training and testing for these databases.

The proposed descriptor performs decently on face recognition, especially for Grimace. That is because the intensity distribution of images keeps stable with variation of expression. Meanwhile the constant face position enables a good recognition result. On the other hand, as variation of lighting and background impacts the intensity of images dramatically, Faces 95 and Faces 96 are burdensome to the descriptor. Fortunately, SVM allows QGFDIs a pleasurable performance against these variations so that a reasonable robustness to lighting and background change is presented.

3.4. Discussions

The experiments on geometric transformation invariance and noise robustness provide supports for QGFDIs that have excellent robustness to geometric transformation and noise. Moreover, it can be observed from the experiments of comparison that QGFDIs make remarkable progress in the accuracy and robustness of color object recognition. For the sake of simplification, the dimension can be briefly reduced by discarding invariants of high frequency, while PCA reduces the dimension of QGFDIs favorably. As a consequence of the vast volume of image database and numerous classes, QFMMIs and QZMIs recognize several objects incorrectly even in the noise-free case. These challenges also cause the performances of QFMMIs and QZMIs corrupting severely when adding noise into the cases. Since the multi-resolution analysis in both radial and circular directions of the object enables QGFDIs for more discriminative power, their CCP attains 100% without noise. Meanwhile, QGFDIs perform far more stably with the noise level growing up. Additionally, the computation of QGFDIs is quite easier than the moment-based descriptors. Thanks to the relatively straightforward computation, QGFDIs circumvent the time-consuming computation of various order moment invariants and achieve a better performance on color object recognition.

The experiment for color image retrieval demonstrates that QGFDIs perform decently for recognition of real scene images and highly textured images. It also indicates that QGFDIs are comparatively robust to the low radial resolution, and with $R=50$, $T=60$ they are sufficient to describe most images. By cooperating with SVM, QGFDIs are equipped with robustness to variation of backgrounds and lighting, and provide a satisfactory results for face recognition.

Table 5
Recognition rates (%) of PASCAL VOC 2007 database with different radial and angular resolutions.

Resolutions	$T=12$	$T=24$	$T=60$	$T=96$	$T=120$
$R=10$	86.85	98.51	99.39	99.69	100.00
$R=20$	87.05	98.71	99.83	99.90	100.00
$R=50$	88.81	98.85	100.00	100.00	100.00
$R=80$	89.25	99.02	100.00	100.00	100.00
$R=100$	91.46	100.00	100.00	100.00	100.00

Table 6
Recognition rates (%) of Texture database with different radial and angular resolutions.

Resolutions	$T=12$	$T=24$	$T=60$	$T=96$	$T=120$
$R=10$	90.34	99.19	99.49	99.86	100.00
$R=20$	92.20	99.53	100.00	100.00	100.00
$R=50$	93.66	99.76	100.00	100.00	100.00
$R=80$	94.92	99.97	100.00	100.00	100.00
$R=100$	95.90	100.00	100.00	100.00	100.00

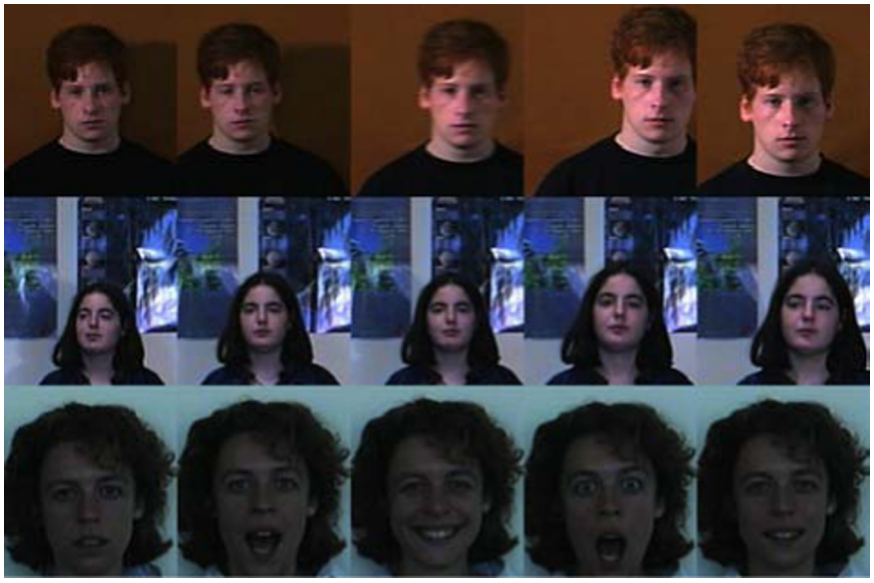

Fig. 8. Examples of each color face database. The first row is from Faces 95, the second row is from Faces 96 and the third row is from Grimace.

Table 7
Results of face recognition using SVM.

Databases	TrPr (%)	TePr (%)	TrRe (%)	TeRe (%)	TrF (%)	TeF (%)
Faces 95	99.96	93.42	99.96	92.01	99.96	91.77
Faces 96	99.96	94.88	99.96	93.58	99.96	93.42
Grimace	99.80	98.98	99.79	98.75	99.79	98.72

In spite of these obvious advantages, there is a limitation on the described approach. The limitation is that by taking the norm on the Fourier spectra to gain rotation invariance, the phase information is discarded. Therefore, an image cannot be reconstructed by the QGFDIs. However, this limitation has no impact on the contribution of our work, as we intend to concentrate on color image description rather than reconstruction.

4. Conclusions

In this paper, we have presented a quaternion generic Fourier descriptor to process color image directly and holistically. Furthermore, the QGFD invariants with respect to geometric transformation have been constructed and applied to color object recognition and color image retrieval. Experimental results demonstrate that the QGFDIs have a remarkable robustness against geometric transformation and noise. Moreover, compared with the moment-based descriptors, QGFDIs offer many benefits in accuracy, robustness and efficiency in color object recognition. In the applications of color image retrieval and face recognition, they perform promisingly as well.

Conflict of interest

The paper should be of interest to readers in the areas of image representation and computer vision.

Acknowledgments

This work is supported by the National Natural Science Foundation of China (61271112). The authors would like to thank Dr. Beijing Chen and Dr. Liqiang Guo for their help on QZM and QFMM.

References

[1] M. Sonka, V. Hlavac, R. Boyle, Image Processing, Analysis, and Machine Vision, Cengage Learning, Inc, Hayward, CA, USA, 2014.
[2] P. Shivakumara, T.Q. Phan, C.L. Tan, New Fourier-statistical features in rgb space for video text detection, IEEE Trans. Circuits Syst. Video Technol. 20 (11) (2010) 1520–1532.
[3] D. Zhang, G. Lu, Study and evaluation of different fourier methods for image retrieval, Image Vis. Comput. 23 (1) (2005) 33–49.
[4] F. Mindru, T. Tuytelaars, L.V. Gool, T. Moons, Moment invariants for recognition under changing viewpoint and illumination, Comput. Vis. Image Underst. 94 (1) (2004) 3–27.
[5] F. Smach, C. Lemaitre, J.-P. Gauthier, J. Miteran, M. Atri, Generalized Fourier descriptors with applications to objects recognition in svm context, J. Math. Imag. Vis. 30 (1) (2008) 43–71.
[6] T. Suk, J. Flusser, Affine moment invariants of color images, in: Proceedings of Computer Analysis of Images and Patterns, Springer, Berlin, Heidelberg, 2009, pp. 334–341.
[7] K.E. Van de Sande, T. Gevers, C.G. Snoek, Evaluating color descriptors for object and scene recognition, IEEE Trans. Pattern Anal. Mach. Intell. 32 (9) (2010) 1582–1596.
[8] T.A. Ell, S.J. Sangwine, Hypercomplex Fourier transforms of color images, IEEE Trans. Image Process. 16 (1) (2007) 22–35.
[9] S.J. Sangwine, Fourier transforms of colour images using quaternion or hypercomplex, numbers, Electron. Lett. 32 (21) (1996) 1979–1980.
[10] S.J. Sangwine, T.A. Ell, Hypercomplex Fourier transforms of color images, in: Proceedings of the 2001 International Conference on Image Processing, ICIP 2001, vol. 1, IEEE, Thessaloniki, Greece, 2001, pp. 137–140.
[11] X. Geng, X. Hu, J. Xiao, Quaternion switching filter for impulse noise reduction in color image, Signal Process. 92 (1) (2012) 150–162.
[12] S. Gai, P. Liu, J. Liu, X. Tang, A new image denoising algorithm via bivariate shrinkage based on quaternion wavelet transform, J. Comput. Inf. Syst. 6 (11) (2010) 3751–3760.
[13] C.E. Moxey, S.J. Sangwine, T.A. Ell, Color–grayscale image registration using hypercomplex phase correlation, in: Proceedings of the 2002 International Conference on Image Processing, ICIP 2002, vol. 3, IEEE, Rochester, New York, USA, 2002, pp. 247–250.
[14] C. Yang, J.-Q. Zhang, X. Xu, H.-H. Chang, G.-J. He, Quaternion phase-correlation-based clutter metric for color images, Opt. Eng. 46 (12) (2007) 127008–127008-7.
[15] B. Chen, G. Coatrieux, G. Chen, X. Sun, J.L. Coatrieux, H. Shu, Full 4-d quaternion discrete fourier transform based watermarking for color images, Digit. Signal Process. 28 (2014) 106–119.
[16] C. Guo, Q. Ma, L. Zhang, Spatio-temporal saliency detection using phase spectrum of quaternion Fourier transform, in: Proceedings of the 2008 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2008, IEEE, Anchorage, Alaska, USA, 2008, pp. 1–8.
[17] C. Guo, L. Zhang, A novel multiresolution spatiotemporal saliency detection model and its applications in image and video compression, IEEE Trans. Image Process. 19 (1) (2010) 185–198.
[18] A. Toet, Computational versus psychophysical bottom-up image saliency: a comparative evaluation study, IEEE Trans. Pattern Anal. Mach. Intell. 33 (11) (2011) 2131–2146.
[19] M.-K. Hu, Visual pattern recognition by moment invariants, IRE Trans. Inf. Theory 8 (2) (1962) 179–187.

[20] M.R. Teague, Image analysis via the general theory of moments, J. Opt. Soc. Am. 70 (8) (1980) 920–930.
[21] Y. Sheng, L. Shen, Orthogonal Fourier–Mellin moments for invariant pattern recognition, J. Opt. Soc. Am. A 11 (6) (1994) 1748–1757.
[22] L.-Q. Guo, M. Zhu, Quaternion Fourier–Mellin moments for color images, Pattern Recognit. 44 (2) (2011) 187–195.
[23] Z. Yang, S.-I. Kamata, Hypercomplex polar Fourier analysis for image representation, IEICE Trans. Inf. Syst. 94 (8) (2011) 1663–1670.
[24] Z. Yang, S.-i. Kamata, Fast hypercomplex polar Fourier analysis, IEICE Trans. Inf. Syst. 95 (4) (2012) 1166–1169.
[25] B. Chen, H. Shu, H. Zhang, G. Chen, C. Toumoulin, J.-L. Dillenseger, L.M. Luo, Quaternion Zernike moments and their invariants for color image analysis and object recognition, Signal Process. 92 (2) (2012) 308–318.
[26] Z. Shao, H. Shu, J. Wu, B. Chen, J.L. Coatrieux, Quaternion Bessel–Fourier moments and their invariant descriptors for object reconstruction and recognition, Pattern Recognit. 47 (2) (2014) 603–611.
[27] D. Zhang, G. Lu, Shape-based image retrieval using generic Fourier descriptor, Signal Process.: Image Commun. 17 (10) (2002) 825–848.
[28] R.B. Yadav, N.K. Nishchal, A.K. Gupta, V.K. Rastogi, Retrieval and classification of shape-based objects using Fourier, generic Fourier, and wavelet-Fourier descriptors technique: a comparative study, Opt. Lasers Eng. 45 (6) (2007) 695–708.
[29] N.S. Vassilieva, Content-based image retrieval methods, Program. Comput. Softw. 35 (3) (2009) 158–180.
[30] D. Zhang, G. Lu, Evaluation of similarity measurement for image retrieval, in: Proceedings of the 2003 International Conference on Neural Networks and Signal Processing, vol. 2, IEEE, Nanjing, China, 2003, pp. 928–931.
[31] S.A. Nene, S.K. Nayar, H. Murase, Columbia object image library (COIL-20), Technical Report CUCS-005-96, Columbia University, New York, 1996.
[32] C.-W. Chong, P. Raveendran, R. Mukundan, A comparative analysis of algorithms for fast computation of Zernike moments, Pattern Recognit. 36 (3) (2003) 731–742.
[33] J.M. Geusebroek, G.J. Burghouts, A.W. Smeulders, The Amsterdam library of object images, Int. J. Comput. Vis. 61 (1) (2005) 103–112.
[34] M. Everingham, L. Van Gool, C. Williams, J. Winn, A. Zisserman, The Pascal Visual Object Classes Challenge 2007 (VOC 2007) Results, Available at: ⟨http://pascallin.ecs.soton.ac.uk/challenges/VOC/voc2007/⟩.
[35] M. Franc, Texture Library, Available at: ⟨http://textures.forrest.cz/⟩.
[36] L. Spacek, Description of the Collection of Facial Images, University of Essex, United Kingdom. Available at: ⟨http://cswww.essex.ac.uk/mv/allfaces/⟩.

Heng Li was born in Zhengzhou, China, in 1989. He received the B.S. degree in communication engineering from Zhengzhou University, China, in 2012. He is currently pursuing Ph.D. degree in Beijing Institute of Technology, China. His main research interests include image processing and medical image analysis.

Zhiwen Liu received his B.S. degree from Xidian University, China, in 1983, and his M.S. and Ph.D. degrees from Beijing Institute of Technology, China, in 1986 and 1989 respectively all in electronic engineering. He is currently a full professor at the School of Information and Electronics, Beijing Institute of Technology. His research interests include array signal processing, biomedical signal and image processing, smart wearable medical devices, estimation and detection theory. He has authored and coauthored over 200 papers in the above areas. He is a senior member of the Chinese Institute of Electronics (CIE) and a member of IEEE.

Yali Huang received the B.S. degree in electronics engineering and the M.S. degree in communication and information system from Hebei University, China, in 2002 and 2005 respectively. She is currently a Ph.D. candidate from Beijing Institute of Technology, China. Her research interests include image processing, active contour models in image segmentation, optical flow and computer vision.

Yonggang Shi received a Ph.D. in electronics engineering from Chinese Academy of Sciences and is currently an associate professor at the Beijing Institute of Technology. His current research interests include medical image segmentation, medical image registration, three dimensional visualization and cell image process and analysis. He has served as a visiting associate professor in the School of Medicine at Stanford University.