

Research Article

Accurate Quaternion Polar Harmonic Transform for Color Image Analysis

Lina Zhang ¹, Yu Sang ², and Donghai Dai ¹

¹Teaching Supervision Department, Zhejiang College of Security Technology, Wenzhou 325016, China

²School of Electronic and Information Engineering, Liaoning Technical University, Huludao 125105, China

Correspondence should be addressed to Lina Zhang; zhangln19802020@163.com

Received 2 November 2021; Accepted 29 November 2021; Published 29 December 2021

Academic Editor: Ahmed Farouk

Copyright © 2021 Lina Zhang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Polar harmonic transforms (PHTs) have been applied in pattern recognition and image analysis. But the current computational framework of PHTs has two main demerits. First, some significant color information may be lost during color image processing in conventional methods because they are based on RGB decomposition or graying. Second, PHTs are influenced by geometric errors and numerical integration errors, which can be seen from image reconstruction errors. This paper presents a novel computational framework of quaternion polar harmonic transforms (QPHTs), namely, accurate QPHTs (AQPHTs). First, to holistically handle color images, quaternion-based PHTs are introduced by using the algebra of quaternions. Second, the Gaussian numerical integration is adopted for geometric and numerical error reduction. When compared with CNNs (convolutional neural networks)-based methods (i.e., VGG16) on the Oxford5K dataset, our AQPHT achieves better performance of scaling invariant representation. Moreover, when evaluated on standard image retrieval benchmarks, our AQPHT using smaller dimension of feature vector achieves comparable results with CNNs-based methods and outperforms the hand craft-based methods by 9.6% w.r.t mAP on the Holidays dataset.

1. Introduction

Rotation-invariant moments (RIMs) are extensively used in image representation and pattern recognition [1–3] because of their outstanding description capability and invariance property. Moreover, RIMs not only provide hand craft features for image representation but also can reconstruct original image from these hand craft features, making them suitable for image watermarking [4–6]. There exist two kinds of RIMs: orthogonal and nonorthogonal. Since the orthogonal RIMs (ORIMs) possess minimum information redundancy and hence better information compactness, they are effective in performance. There are several low order moments and transforms can sufficiently extract the essential features of images. The most popular among these ORIMs [7, 8] are Zernike moments (ZMs) and pseudo ZMs (PZMs). Yap et al. [9] recently introduced a few ORITs, collectively known as PHTs, including the polar sine transforms (PSTs), polar cosine transforms (PCTs), and

polar complex exponential transforms (PCETs). ORIMs are different from ORITs in the radial parts of kernel functions, which are polynomials and sinusoidal functions in ORIMs and ORITs, respectively. Compared with ORIMs, PHTs are more efficient in computation [10]. Besides, the high order transforms of PHTs are numerically stable, whereas those of ORIMs are numerically unstable. Therefore, PHTs are preferred to ORIMs and have recently been utilized in numerous image processing applications.

Especially for small size images, the current computational framework of PHTs is disadvantaged by geometric errors and numerical integration errors, which can be seen from image reconstruction errors, which are visible near the circular disk center, and even low orders of transform are numerically unstable. Mapping a square image into a unit circular disk can cause geometric errors, and approximating the integration with zeroth-order summation will lead to numerical integration errors. Accurate PHTs are extremely essential in tremendous

image processing applications, primarily in template matching and optical character recognition problems using small size images. In addition to the problem of computation accuracy, another problem is that ORIMs and ORITs mainly focus on grayscale images, yet with the continuous improvement of computer performance, color images are drawing more and more attention from researchers in the related fields, because color images can provide much more abundant information compared with grayscale images. Most of the current research on color image moments depends on the intensity or a single channel within the color space of color images, where the information and relationship between color components in a specific color space are discarded. Several traditional orthogonal and nonorthogonal moments have been developed into quaternion moments, e.g., quaternion Zernike moments (QZMs) [2], quaternion radial harmonic Fourier moments (QRHFM) [11], quaternion polar harmonic Fourier moments (QPHFMs) [12], and quaternion polar harmonic transforms (QPHTs) [13, 14]. Although various quaternion image orthogonal moments have been proposed by researchers, most of them give unsatisfied image reconstruction performance. In the present work, accurate quaternion orthogonal transforms, i.e., accurate QPHTs (AQPHTs), are proposed. In summary, the innovations of the proposed AQPHTs include the following: firstly, by using the algebra of quaternions, we extend the PHTs to color images. Secondly, a computational framework, which reduces geometric errors and numerical integration errors, is developed to construct an approach to accurately calculate PHTs. Experiments are conducted to comparatively study the performance in image reconstruction and image retrieval of AQPHT and other ORIMs, and the results present that AQPHT has the best image reconstruction performance and superb behavior in invariant image representation with and without noises.

The paper is organized as follows. Section 2 introduces the quaternion algebra and PHTs. Section 3 describes the computational framework of our AQPHTs. In Section 4, the effectiveness of the proposed AQPHTs is evaluated. Section 5 concludes this study.

2. Preliminaries

2.1. Quaternion Algebra. A quaternion can be regarded as a generalized complex number. The formula of a quaternion is as follows:

$$q = a + bi + cj + dk, \quad (1)$$

where q denotes the quaternion, a and $bi + cj + dk$ denote the real and the imaginary part; a , b , c , and d are real numbers, and i , j , and k are complex operators that satisfy these conditions:

$$\begin{aligned} i^2 &= j^2 = k^2 = ijk = -1, \\ ij &= k, \\ jk &= i, \\ ki &= j, \\ ji &= -k, \\ kj &= -i, \\ ik &= -j. \end{aligned} \quad (2)$$

The conjugate of μ is $\bar{\mu} = a - bi - cj - dk$, and the magnitude equals $|\mu| = \sqrt{a^2 + b^2 + c^2 + d^2}$. A quaternion $\mu = s(\mu) + v(\mu)$ can be considered as the combination of a scalar part $s(\mu) = a$ and a vector part $v(\mu) = bi + cj + dk$.

2.2. Polar Harmonic Transform. The two-dimensional polar harmonic transforms $M_{n,l}$ of an n -order image function $f(r, \theta)$, with repetition l ($|l| \geq 0$) defined by

$$M_{n,l} = \Omega_n \int_0^{2\pi} \int_0^1 H_{n,l}^* f(r, \theta) r dr d\theta, \quad (3)$$

where Ω_n denotes a normalization factor, and $H_{n,l}^*(r, \theta)$ denotes the complex conjugate of the basis function $H_{n,l}(r, \theta)$. For PCET, $\Omega_n = 1/\pi$ and $H_{n,l}(r, \theta)$ are given by

$$H_{n,l}^{\text{PCET}}(r, \theta) = e^{i2\pi nr^2} e^{il\theta}, \quad (4)$$

where $i = \sqrt{-1}$. While in PCT and PST,

$$\Omega_n = \begin{cases} \frac{1}{\pi}, & n = 0, \\ \frac{2}{\pi}, & n \neq 0. \end{cases} \quad (5)$$

The basis functions of PST and PCT are defined, respectively, as

$$\begin{aligned} H_{n,l}^{\text{PST}}(r, \theta) &= \sin(\pi nr^2) e^{il\theta}, \\ H_{n,l}^{\text{PCT}}(r, \theta) &= \cos(\pi nr^2) e^{il\theta}. \end{aligned} \quad (6)$$

The total number of PCET for $n = n_{\max}$ and $l = l_{\max}$ is $(1 + 2n_{\max})(1 + 2l_{\max})$ where $n_{\max} = l_{\max}$. However, PST and PCT have a smaller number of features, which are $n_{\max}(1 + 2l_{\max})$ and $(1 + n_{\max})(1 + 2l_{\max})$, respectively.

For an $N \times N$ image, there is no analytical solution to the integration in equation (3), therefore its zeroth-order approximation is commonly used and given by

$$M_{n,l} = \Omega_n \sum_j \sum_k^{N-1} f(x_j, y_k) H_{n,l}^*(x_j, y_k) \Delta x_j \Delta y_k, \quad x_j^2 + y_k^2 \leq 1, \quad (7)$$

where

$$\begin{aligned} x_j &= \frac{2j+1-N}{N}, \\ y_k &= \frac{2k+1-N}{N}, \quad j, k = 0, 1, \dots, N-1. \end{aligned} \quad (8)$$

We perform inverse transforms to reconstruct the image function. It is expressed by

$$\hat{f}(x_j, y_k) = \sum_{n=n_{\min}}^{n_{\max}} \sum_{l=l_{\min}}^{l_{\max}} M_{n,l} H_{n,l}(x_j, y_k), \quad j, k = 0, 1, \dots, N-1, \quad (9)$$

where n_{\max} , n_{\min} , l_{\max} , and l_{\min} are the maximum and minimum orders and repetitions of PHTs, respectively, as specified in equation (3).

3. Accurate QPHTs

3.1. QPHTs. In traditional processing methods, color images are usually divided into 3 components and each component is separately handled by a follow-up process without taking into account the association within components. But, quaternion-based theory treats a color image as an integral vector, by which the relationships between each component are reflected. Let $f_R(r, \theta)$, $f_G(r, \theta)$, and $f_B(r, \theta)$ denote the color components of red, green, and blue, respectively, then we encode a color pixel $f(r, \theta)$. The pixel can be encoded as a pure quaternion:

$$f(r, \theta) = f_R(r, \theta)i + f_G(r, \theta)j + f_B(r, \theta)k. \quad (10)$$

Therefore, PHTs can be defined in the quaternion field. The multidimension feature of color images can be addressed by the transforms wherein each color pixel is treated as a whole. Due to the noncommutative multiplication of quaternions, each quaternion transform has two different forms. The left-side quaternion PCET is defined as follows:

$$\begin{aligned} M_{n,l}^L &= \frac{1}{\pi} \int_0^{2\pi} \int_0^1 H_{n,l}^* f(r, \theta) r dr d\theta \\ &= \frac{1}{\pi} \int_0^{2\pi} \int_0^1 e^{-\mu l \theta} e^{-\mu 2\pi n r^2} f(r, \theta) r dr d\theta, \end{aligned} \quad (11)$$

where μ denotes a pure unit quaternion and $f(r, \theta)$ denotes the quaternion representation of a color pixel. $\mu = (i + j + k)/\sqrt{3}$, i.e., the gray line in the RGB space is chosen in this work. By reversing the orders of images and transform kernels in equation (11), we can obtain the right-side QPCET:

$$M_{n,l}^R = \frac{1}{\pi} \int_0^{2\pi} \int_0^1 f(r, \theta) e^{-\mu l \theta} e^{-\mu 2\pi n r^2} r dr d\theta. \quad (12)$$

The reconstruction of a color image by using the QPCET coefficients of the left and right sides can be written as equations (13) and (14), respectively:

$$f(r, \theta) = \sum_{n=0}^{n_{\max}} \sum_{l=-l_{\max}}^{l_{\max}} e^{\mu 2\pi n r^2} e^{\mu l \theta} M_{n,l}^L, \quad (13)$$

$$f(r, \theta) = \sum_{n=0}^{n_{\max}} \sum_{l=-l_{\max}}^{l_{\max}} M_{n,l}^R e^{\mu 2\pi n r^2} e^{\mu l \theta}. \quad (14)$$

Analogous to QPCET, the left and right-side QPCT can be defined as equations (15) and (16), respectively,

$$M_{n,l}^L = \Omega_n \int_0^{2\pi} \int_0^1 \cos(\pi n r^2) e^{-\mu l \theta} f(r, \theta) r dr d\theta, \quad (15)$$

$$M_{n,l}^R = \Omega_n \int_0^{2\pi} \int_0^1 f(r, \theta) \cos(\pi n r^2) e^{-\mu l \theta} r dr d\theta, \quad (16)$$

where all factors are defined as in equation (5). Replace the radial components in Equations (15) and (16) with $f_B(r, \theta)$, and we can obtain the QPST. The construction of color images can also be achieved by using QPCT and QPST coefficients.

We note that the QPHT in this paper refers to the left-side QPCET $M_{n,l}^L$ as defined in equation (11), unless otherwise specified.

3.2. Computation of Accurate QPHT. This subsection provides a computational framework for calculating QPHTs, which can reduce geometric errors and numerical integration errors. Using our method, only the part inside the unit disk of the pixel is considered. For numerical integration, we rewrite equation (11) as

$$M_{n,l}^L = \frac{1}{\pi} \sum_{j=0}^{N-1} \sum_{k=0}^{N-1} f(x_j, y_k) h_{nl,jk}, \quad x_j^2 + y_k^2 \leq 1. \quad (17)$$

We assume the image function $f_B(r, \theta)$ is constant in one grid. The calculation accuracy of QPHTs can be increased through computing the numerical integration of basis function $h_{nl,jk}$ for each grid, that is,

$$h_{nl,jk} = \int_{a_j}^{a_{j+1}} \int_{b_k}^{b_{k+1}} H_{nl}^*(x, y) dx dy, \quad a_j^2 + b_k^2 \leq 1, \quad (18)$$

where

$$\begin{aligned} a_j &= \frac{2j-N}{N}, \\ b_k &= \frac{2k-N}{N}, \quad j, k = 0, 1, \dots, N. \end{aligned} \quad (19)$$

Since equation (18) is evaluated by using a $g \times g$ point Gaussian numerical integration method, it can be simplified into

$$h_{nl,jk} \cong \frac{1}{N^2} \sum_{l=0}^{g-1} \sum_{m=0}^{g-1} \omega_l \omega_m \times H_{nl}^* \left(\frac{t_l + 2j + 1 - N}{N}, \frac{t_m + 2k + 1 - N}{N} \right) \cdot \left(\frac{t_l + 2j + 1 - N}{N} \right)^2 + \left(\frac{t_m + 2k + 1 - N}{N} \right)^2 \leq 1. \quad (20)$$

The values of ω and t for a given g can be obtained by using standard procedures [15]. For a quick reference, we provide these values in Table 1 for $g=1$ through 8. The constraint in equation (21) is an improvement over the constraint $x_j^2 + y_k^2 \leq 1$ in the zeroth-order approximation.

If the center of a grid falls outside the unit circle, this grid will be completely ignored in computation in zeroth-order approximation. But in the new constraint, a grid will be considered in the calculation if its sampling point falls within the unit disk even if its centers fall outside the unit disk. An improvement of the approximation of the unit disk has been observed with this new constraint. This is a great contribution of the numerical integration method for reducing geometric errors and numerical integration errors simultaneously by ensuring that none of the sampling points fall outside the unit disk. We find that the performance of image reconstruction improves as we increase the number of g , but the gain starts to saturate at $g=6$. Therefore, the accurate computation experiment of QPHTs uses 5×5 sampling points (i.e., $g=5$), which we take as a tradeoff accuracy and speed for color images.

3.3. Geometric Invariance of AQPHT. Here, we will derive and analyze the rotation and scaling invariant property of AQPHTs.

3.3.1. Rotation Invariance. Let $\hat{f}(r, \theta) = f(r, \theta + \alpha)$ denote the image $f(r, \theta)$ rotated by the angle α . Accordingly, the left-side AQPHT of $\hat{f}(r, \theta)$ is

$$\begin{aligned} \hat{M}_{n,l}^L &= \frac{1}{2\pi} \int_0^{2\pi} \int_0^1 \hat{f}(r, \theta) \exp(-\mu 2\pi n r^2) \exp(-\mu l \theta) r dr d\theta \\ &= \frac{1}{2\pi} \int_0^{2\pi} \int_0^1 f(r, \theta + \alpha) \exp(-\mu 2\pi n r^2) \exp(-\mu l \theta) r dr d\theta \frac{1}{2\pi} \\ &= \int_0^{2\pi} \int_0^1 f(r, \theta) \exp(-\mu 2\pi n r^2) \exp(-\mu l \theta) r dr d\theta \times \exp(\mu l \theta) \\ &= M_{n, \times \exp(\mu l \theta), l}, \end{aligned} \quad (21)$$

where $M_{n,l}$ and $\hat{M}_{n,l}$ are the AQPHTs of $f(r, \theta)$ and $\hat{f}(r, \theta)$, respectively.

Accordingly to equation (21), we know that a rotation of the color image by an angle α induces a phase shift $e^{\mu l \theta}$ of the $M_{n,l}^L(f)$. Taking the norm on both sides of equation (17), we have

TABLE 1: Weights and locations of sampling points for $g \times g$ point integration.

g	ω	t	g	ω	t
1	2.0	0.0	7	0.12948497	± 0.94910791
2	1.0	± 0.57735027		0.27970539	± 0.74153119
3	0.55555556	± 0.77459667		0.38183005	± 0.40584515
	0.88888889	0.0		0.41795918	0.0
4	0.34785485	± 0.86113631	8	0.10122853	± 0.96028987
	0.65214515	± 0.33998104		0.22238103	± 0.79666648
0.23692688	± 0.90617985	0.31370665		± 0.52553241	
5	0.47862867	± 0.53846931		0.36268378	± 0.18343464
	0.56888889	0.0			
	0.17132449	± 0.93246951			
6	0.36076157	± 0.66120939			
	0.46791393	± 0.23861919			

$$\begin{aligned} |\hat{M}_{n,l}| &= |M_{n,l} \exp(\mu l \theta)| \\ &= |M_{n,l}| |\exp(\mu l \theta)| \\ &= |M_{n,l}|. \end{aligned} \quad (22)$$

Therefore, the rotation invariant can be achieved by taking the norm of AQPHTs. In other words, the AQPHTs modulus coefficients $|M_{n,l}^L(f)|$ are invariant with respect to image rotation.

3.3.2. Scaling Invariance. Theoretically, AQPHTs are not invariant to image scaling, but scaling invariance can be obtained by normalizing the image into a unit circle. If an image $f(r, \theta)$ with $N \times N$ pixels is mapped to a unit circle $(x_j, y_k) \in [-1, 1] \times [-1, 1]$ with $x_j = 2j - N + 1/N$, $y_k = N - 1 - 2k/N$, ($j, k = 0, 1, \dots, N - 1$) and a unit circle is made to cover the same contents of the image, the AQPHTs are variant to image scaling.

4. Experiments and Analysis

This section is intended to validate the effectiveness of AQPHT invariants for color images. The experiments are performed using MATLAB version 8.6 on a 2.9 GHz processor, 8 GB RAM computer, and Microsoft Windows 10 Ultimate operating system.

4.1. Experiment on Scaling Invariance. Numerous great classification models, e.g., VGG [16] and ResNet [17], have been developed on the basis of the powerful deep learning framework. Their recognition accuracy is amazing and even higher than that of human beings. However, the performance is achieved only for large size images having rich object structure and high quality appearance. The resolution of images with small size is low, which limits the learning of discriminative representations, thus leading to identification failure [18]. In this experiment, the impact of down-scaling operation on image representations can be evaluated on Oxford5K dataset [19] by our AQPHT and the widespread deep model of VGG16. For convenience, only the neural activation in the 36-th layers of the VGG16 model is studied as an example of deep models in this work. The ability of

representation is measured by the mean Euclidean distance which is calculated between the features extracted from an image at various resolutions. Figure 1 separately presents the experimental results of mean Euclidean distance at various down-sampling scales for the VGG16 model and AQPHT. From Figure 1(a), although the VGG16 model is carefully trained with some extra useful training tricks on the Oxford5K dataset, the mean Euclidean distance is growing larger with the decrease of image resolution, indicating that the ability of deep representation reduces with decreasing image resolution. From Figure 1(b), the mean Euclidean distance of the AQPHT is almost constant, demonstrating the superb performance of the AQPHT invariants under scaling transform.

4.2. Experiment on Image Reconstruction. The capability of image representation of image moments can be measured by image reconstruction performance. This subsection compares the performance of image reconstruction between AQPHT and QZM [2], QPHT [14], QPHFM [12], and QRHFM [11]. Following the standard [11, 12], we set the relevant parameters in our experiments. The mean square reconstruction error (MSRE) is used to measure the reconstruction performance [11, 12]; lower MSREs equate to better reconstruction performance. Suppose $f(x, y)$ and $\hat{f}(x, y)$ indicate the original image and reconstructed image, respectively; the MSRE is defined as

$$\text{MSRE} = \frac{\sum_{x=0}^{N-1} \sum_{y=0}^{N-1} |f(x, y) - \hat{f}(x, y)|^2}{\sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f^2(x, y)}. \quad (23)$$

In Figure 2, we plot the curves of average MSREs versus the data amount of coefficients. The coefficients in image reconstruction can be represented by real values, and the total number of required real values is defined as data amount. For instance, four real values are needed to represent one AQPHT coefficient, while representing the coefficients of three components of APHT-RGB require six real values [9]. It can be seen from Figure 2 that our AQPHT achieves higher MSREs than APHT-RGB. By dealing with various color channels in a holistic way, AQPHTs can obtain the essence of color images from the inter- and intrachannel directions, demonstrating better compactness.

Next, we use two additional metrics, i.e., PSNR and SSIM, to evaluate the reconstruction performance. The PSNR between the reconstructed image \hat{Y} and original image Y of size is calculated as

$$\text{PSNR}(\hat{Y}, Y) = 10 \log_{10} \frac{\sum_{x=0}^{N-1} \sum_{y=0}^{N-1} \text{MAX}_I^2}{\sum_{x=0}^{N-1} \sum_{y=0}^{N-1} (\hat{Y}(x, y) - Y(x, y))^2}, \quad (24)$$

where MAX_I is the maximum possible pixel intensity value. The SSIM is designed to better match the human perception compared to PSNR; the SSIM is defined as

$$\text{SSIM}(\hat{Y}, Y) = \frac{(2\mu_{\hat{Y}}\mu_Y + c_1)(2\sigma_{\hat{Y}Y} + c_2)}{(\mu_{\hat{Y}}^2 + \mu_Y^2 + c_1) + (\sigma_{\hat{Y}}^2 + \sigma_Y^2 + c_2)}, \quad (25)$$

where μ and σ are respectively the average and variance of the pixel values, $\sigma_{\hat{Y}Y}$ is the covariance of \hat{Y} and Y ; c_1 and c_2 are constants. Following previous methods [20, 21], we set $c_1 = 0.01$ and $c_2 = 0.03$ in this paper. Higher values for PSNR and SSIM indicate better performance.

The ‘‘Lena’’ of size of 128×128 is reconstructed using our AQPHT and the compared methods with a maximum order ranging from 4 to 28. As shown in Figure 3, we can find that the reconstructed images of our AQPHT are far better than other moments. When the number of moments exceeds a certain value, the reconstruction performance of QZM [2] and QPHT [14] even degrades, whereas the reconstruction performance of our proposed AQPHT keeps getting better.

4.3. Experiment on Image Retrieval. Here, a series of extensive experiments are conducted to compare our method with other leading edge ones, namely, traditional hand craft-based methods [22–25] and CNN-based methods [26–30]. To evaluate the performance, we use the average precision (AP) measure computed as the area under the precision-recall curve for a query. We compute an average precision score for each of the 5 queries for a landmark, averaging these to obtain a mean average precision (mAP) score. The average of these mAP scores is used as a single number to evaluate the overall performance. For a fair comparison, postprocessing methods, e.g., query expansion are excluded in the work, and only mAP in representation with relevant feature dimensions is reported. The results of retrieval accuracy (mAP) [31] of the UK-bench [32], Holidays [33], and Oxford5k [19] are presented in Table 2, in which the bold indicates the best results. From Table 2, the performance of our method is much better than all hand craft-based methods but marginally worse than some CNN-based methods [28–30]. However, the length of a feature vector largely determines its retrieval efficiency. We note that

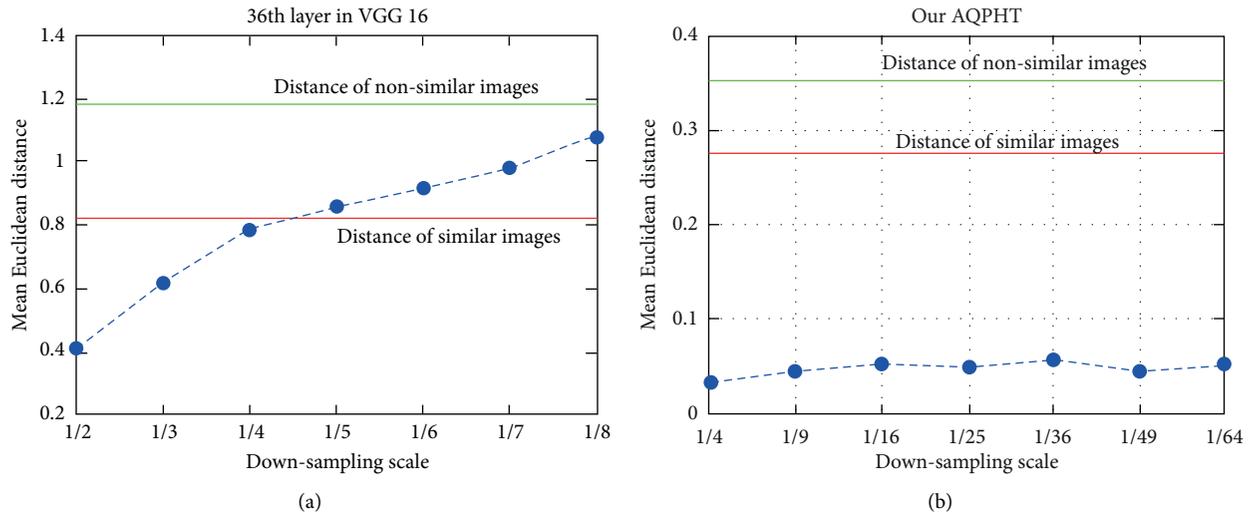


FIGURE 1: Mean Euclidean distance at various low resolutions on Oxford5K dataset. (a) VGG16 and (b) the proposed AQPHT.

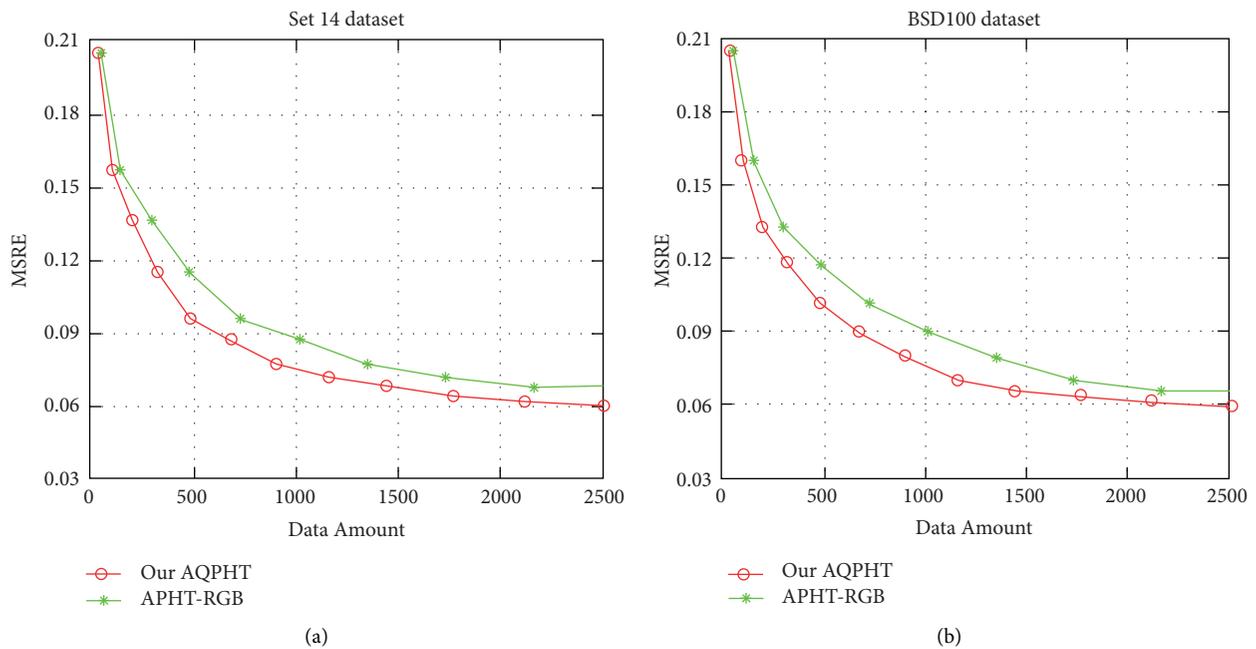


FIGURE 2: The MSREs curves of APH-RGB and our AQPHT on (a) Set 14 and (b) BSD 100.

our method uses feature vectors with significantly low dimensions comparing with CNN-based methods. In other words, the retrieval efficiency of our method is increased under the same experimental conditions in

spite of the slight decrease of its accuracy, which means our method is a more effective compromising method for color image retrieval. Besides, the proposed AQPHT performs better than the QPHT [31] on 3 datasets by a

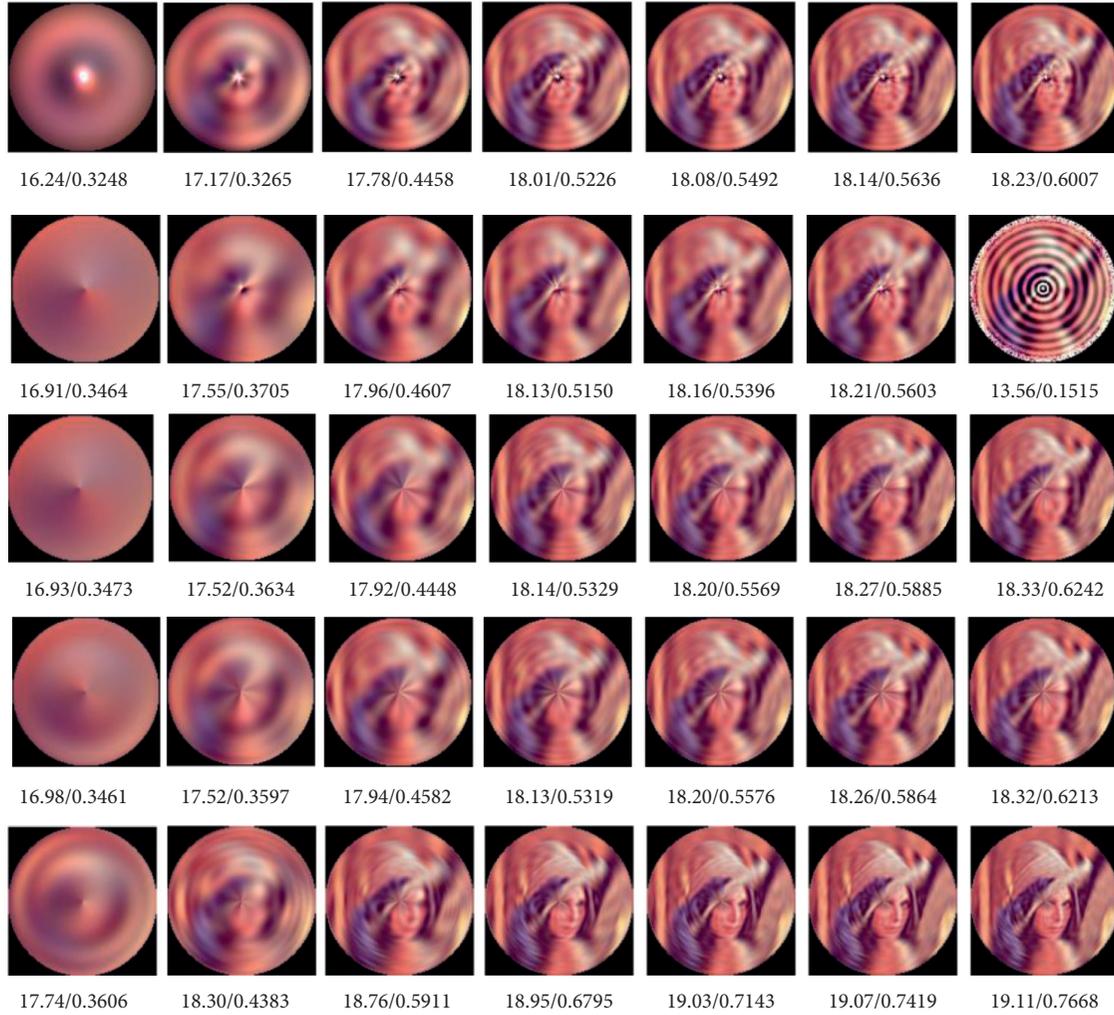


FIGURE 3: Comparison of reconstructed images for Lena, w.r.t. PSNR and SSIM. From top to down are QZM, QPHT, QPHFM, QRHFM, and our AQPHT, respectively.

TABLE 2: Comparison of retrieval performance in terms of mean average precision (mAP).

Method	Feature vector length	Oxford5K	Holidays	UK-bench	
Hand craft-based methods	LCS + RN [22]	0.517	0.658	0.324	
	Improved BOW [23]	0.546	0.747	0.352	
	LTF [24]	—	0.689	0.355	
	LBP + color [25]	0.664	0.735	0.347	
CNN-based methods	Neural codes [26]	128	0.557	0.789	0.356
	CRB-M [27]	512	0.582	0.809	0.340
	VGG + GeM [28]	512	0.820	0.831	—
	AWF [29]	512	0.811	0.826	0.360
	Triplet network [30]	512	0.820	0.842	0.362
QPHT [31]	95	0.696	0.812	0.356	
Our AQPHT	95	0.742	0.831	0.363	

The bold values represent results of our proposed method, which denote the best results among the methods in Table 2.

large margin, which demonstrates the image representation ability of AQPHT which is stronger than that of QPHT.

5. Conclusion

This paper presents a novel computational framework of quaternion-based polar harmonic transform, namely, accuracy quaternion polar harmonic transforms (AQPHTs). Firstly, to holistically deal with color images, AQPHTs are introduced based on the algebra of quaternions. Secondly, geometric errors and numerical errors are reduced by using the Gaussian numerical integration. Many comparative experiments are conducted to analyze the performance of AQPHTs and other ORIMs. Experimental results verify the superb performance of AQPHTs in image reconstruction and invariant image representation. For future work, AQPHTs would be tested in other color image processing domains, e.g., watermarking, segmentation, and retrieval. Besides, more accurate algorithms will be put forward and the computational method of quaternion moments will be improved.

Data Availability

The data included in this paper are available without any restriction from the corresponding author.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This work was supported in part by the National Science Foundation of China (NSFC) under Grant no. 61602226 and in part by the PhD Startup Foundation of Liaoning Technical University of China under Grant no. 18-1021.

References

- [1] B. Aneta, K. Przemyslaw, and S. Dariusz, "Constant-time calculation of Zernike moments for detection with rotational invariance," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 41, no. 3, pp. 537–551, 2019.
- [2] H. Y. Yang, S. R. Qi, C. Wang, S. B. Yang, and X. Y. Wang, "Image analysis by log-polar exponent-fourier moments," *Pattern Recognition*, vol. 101, Article ID 107177, 2020.
- [3] K. Arindam, P. Sourav, C. Arghya, B. Debotosh, S. L. H. Edmond, and P. H. S. Hubert, "LMZMPM: local modified Zernike moment per-unit mass for robust human face recognition," *IEEE Transactions on Information Forensics and Security*, vol. 16, pp. 495–509, 2021.
- [4] C. Wang, X. Wang, Z. Xia, and C. Zhang, "Ternary radial harmonic fourier moments based robust stereo image zero-watermarking algorithm," *Information Sciences*, vol. 470, pp. 109–120, 2019.
- [5] Z. Xia, X. Wang, W. Zhou, R. Li, C. Wang, and C. Zhang, "Color medical image lossless watermarking using chaotic system and accurate quaternion polar harmonic transforms," *Signal Processing*, vol. 157, pp. 108–118, 2019.
- [6] Z. Xia, X. Wang, C. Wang, B. Ma, M. Wang, and Y. Q. Shi, "Local quaternion polar harmonic Fourier moments-based multiple zero-watermarking scheme for color medical images," *Knowledge-Based Systems*, vol. 216, Article ID 106568, 2021.
- [7] A. W. Deng, C. H. Wei, and C. Y. Gwo, "Stable, fast computation of high-order Zernike moments using a recursive method," *Pattern Recognition*, vol. 56, pp. 16–25, 2016.
- [8] Y. D. Zhang, Y. Jiang, W. Zhu, S. Lu, and G. Zhao, "Exploring a smart pathological brain detection method on pseudo Zernike moment," *Multimedia Tools and Applications*, vol. 77, no. 17, pp. 22589–22604, 2018.
- [9] P. T. Yap, X. Jiang, and A. C. Kot, "Two-dimensional polar harmonic transforms for invariant image representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 7, pp. 1259–1270, 2010.
- [10] K. M. Hosny and M. M. Darwish, "Highly accurate and numerically stable higher order QPCET moments for color image representation," *Pattern Recognition Letters*, vol. 97, pp. 29–36, 2017.
- [11] Y. Liu, S. Zhang, G. Li, H. Wang, and J. Yang, "Accurate quaternion radial harmonic fourier moments for color image reconstruction and object recognition," *Pattern Analysis & Applications*, vol. 23, no. 4, pp. 1551–1567, 2020.
- [12] C. Wang, X. Wang, Y. Li, Z. Xia, and C. Zhang, "Quaternion polar harmonic fourier moments for color images," *Information Sciences*, vol. 450, pp. 141–156, 2018.
- [13] Y. N. Li, "Quaternion polar harmonic transforms for color images," *IEEE Signal Processing Letters*, vol. 20, no. 8, pp. 803–806, 2013.
- [14] Y. Liu, S. Zhang, and J. Yang, "Color image watermark decoder by modeling quaternion polar harmonic transform with BKF distribution," *Signal Processing: Image Communication*, vol. 88, Article ID 115946, 2020.
- [15] C. Singh, S. Pooja, and R. Upneja, "On image reconstruction, numerical stability, and invariance of orthogonal radial moments and radial harmonic transforms," *Pattern Recognition and Image Analysis*, vol. 21, no. 4, pp. 663–676, 2011.
- [16] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," in *Proceedings of the International Conference on Learning Representations*, San Diego, CA, USA, May 2015.
- [17] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 770–778, CVPR), Las Vegas, NV, USA, December 2016.
- [18] W. Tan, B. Yan, and B. Bare, "Feature super-resolution: make machine see more clearly," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3994–4002, CVPR), Salt Lake City, UT, USA, December 2018.
- [19] J. Philbin, O. Chum, M. Isard, J. Sivic, and A. Zisserman, "Object retrieval with large vocabularies and fast spatial matching," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1–8, CVPR), Minneapolis, MN, USA, July 2007.
- [20] Y. Liu, S. Zhang, J. Xu, J. Yang, and Y.-W. Tai, "An accurate and lightweight method for human body image super-resolution," *IEEE Transactions on Image Processing*, vol. 30, pp. 2888–2897, 2021.
- [21] Y. Liu, S. Zhang, C. Wang, and J. Xu, "Single image super-resolution via hybrid resolution NSST prediction," *Computer Vision and Image Understanding*, vol. 207, Article ID 103202, 2021.

- [22] J. Delhumeau, P. H. Gosselin, H. Jégou, and P. Pérez, “Revisiting the VLAD image representation,” in *Proceedings of the 21st ACM International Conference on Multimedia*, pp. 653–656, Barcelona, Spain, October 2013.
- [23] H. Jégou, M. Douze, and C. Schmid, “Improving bag-of-features for large scale image search,” *International Journal of Computer Vision*, vol. 87, no. 3, pp. 316–336, 2010.
- [24] S. K. Kanaparthi, U. S. N. Raju, P. Shanmukhi, G. Khyathi Aneesha, and M. E. U. Rahman, “Image retrieval by integrating global correlation of color and intensity histograms with local texture features,” *Multimedia Tools and Applications*, vol. 79, no. 47, pp. 34875–34911, 2020.
- [25] M. O. Divya and E. R. Vimina, “Content based image retrieval with multi-channel LBP and colour features,” *International Journal of Applied Pattern Recognition*, vol. 6, no. 2, pp. 177–193, 2020.
- [26] A. Babenko, A. Slesarev, A. Chigorin, and V. Lempitsky, “Neural codes for image retrieval,” in *Proceedings of the European Conference on Computer Vision*, pp. 584–599, ECCV), Zurich, Switzerland, September 2014.
- [27] A. Alzu’bi, A. Amira, and N. Ramzan, “Content-based image retrieval with compact deep convolutional features,” *Neuro-computing*, vol. 249, pp. 95–105, 2017.
- [28] F. Radenović, G. Tolias, and O. Chum, “Fine-tuning CNN image retrieval with No human annotation,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 41, no. 7, pp. 1655–1688, 2018.
- [29] J. Qin, J. Chen, X. Xiang, Y. Tan, W. Ma, and J. Wang, “A privacy-preserving image retrieval method based on deep learning and adaptive weighted fusion,” *Journal of Real-Time Image Processing*, vol. 17, no. 1, pp. 161–173, 2020.
- [30] W. Min, S. Mei, Z. Li, and S. Jiang, “A two-stage triplet network training framework for image retrieval,” *IEEE Transactions on Multimedia*, vol. 22, no. 12, pp. 3128–3138, 2020.
- [31] Y. N. Liu, S. S. Zhang, Y. Sang, and S. M. Wang, “Improving image retrieval by integrating shape and texture features,” *Multimedia Tools and Applications*, vol. 78, no. 2, pp. 2525–2550, 2019.
- [32] D. Nister and H. Stewenius, “Scalable recognition with a vocabulary tree,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2161–2168, New York, NY, USA, October 2006.
- [33] H. Jegou, M. Douze, and C. Schmid, “Hamming embedding and weak geometric consistency for large scale image search,” in *Proceedings of the European Conference on Computer Vision*, pp. 304–317, Marseille, France, October 2008.