IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received March 1, 2019, accepted March 22, 2019, date of publication March 28, 2019, date of current version April 16, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2908029

Perceptual Image Hashing Based on Weber Local Binary Pattern and Color Angle Representation

CHUAN QIN[1,2], YECEN HU[1], HENG YAO[1], (Member, IEEE), XINTAO DUAN[3], AND LIPING GAO[1,2]

[1]School of Optical-Electrical and Computer Engineering, University of Shanghai for Science and Technology, Shanghai 200093, China
[2]Shanghai Key Laboratory of Data Science, Fudan University, Shanghai 200433, China
[3]College of Computer and Information Engineering, Henan Normal University, Henan 453007, China

Corresponding author: Heng Yao (hyao@usst.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 61672354 and Grant 61702332.

ABSTRACT This paper proposes an efficient scheme for generating image hashing by combining the local texture and color angle features. During the stage of texture extraction, using Weber's Law, the difference ratios between the center pixels and their surrounding pixels are calculated and the dimensions of these values are further reduced by applying principal component analysis to the statistical histogram. In the stage of color feature extraction, the color angle of each pixel is computed before dimensional reduction and is carried out using a discrete cosine transform and a significant coefficients selection strategy. The main contribution of this paper is a novel construction for image hashing that incorporates texture and color features by using Weber local binary pattern and color angular pattern. The experimental results demonstrate the efficacy of the proposed scheme, especially for the perceptual robustness against common content-preserving manipulations, such as the JPEG compression, Gaussian low-pass filtering, and image scaling. Based on the comparisons with the state-of-the-art schemes, receiver operating characteristic curves and integrated histograms of normalized distances show the superiority of our scheme in terms of robustness and discrimination.

INDEX TERMS Image hashing, Weber's law, local binary pattern, color angular pattern.

I. INTRODUCTION

With the development of multimedia technology, human society has entered into an information age that offers convenience in the capturing and transmitting of digital images. Since digital images contain large amounts of data, have high redundancy and low confidentiality, and are insensitive to distortion and easy to edit, the phenomena of digital image replication and malicious tampering is relatively serious, as it can lead to a copyright authentication problem. Although several digital watermarking methods [1], [2], [3] have been proposed to verify the authentication of images, there are two main issues with this kind of technique: firstly, it causes some degree of image distortion during the process of embedding of a watermark, and secondly, it is necessary to consider the capability of watermark extraction. In addition, due to the increase in the amount of image data, the retrieval of certain required images from a high-volume image database with satisfactory efficiency is another important task for image retrieval applications [4].

An image hashing technique is proposed and developed here to solve these problems. This involves a short, fixed-length sequence of numbers that can be used to represent the features of an image. Since image hashing is a compressed expression based on image visual content, like a fingerprint of an image, it is also referred to as an image fingerprint. The generation of an image hash does not change the content of the image, and the small amount of data involved greatly facilitates the retrieval requirements, making it suitable for both the authentication and retrieval of images. In addition, image hashing can be applied in tamper detection, digital watermarking, and other areas. In cryptography, hash functions such as MD5 and SHA-1 are sensitive to changes; that is to say, any change in input data results in a completely different output. Unlike the hash function in cryptography, image hashing has the following properties. *(1) Perceptual robustness:* the image may undergo conventional digital processing such as enhancement, filtering, noise interference or

The associate editor coordinating the review of this manuscript and approving it for publication was Orazio Gambino.

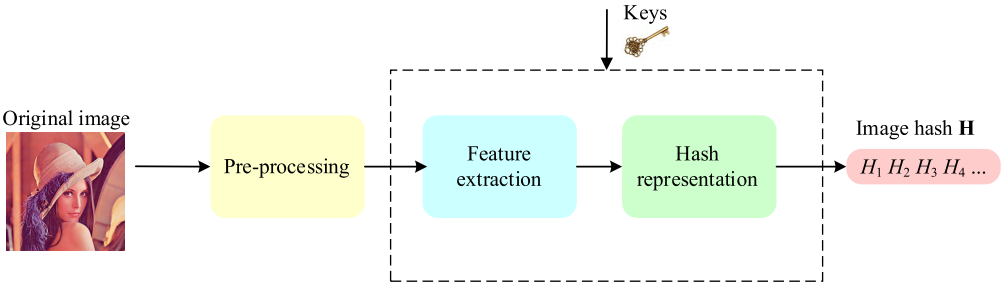

FIGURE 1. Block diagram of image hashing.

geometrical transforms. Although its internal data structure changes slightly, its visual content largely remains the same; hence, the similarity of two image hashes should be kept at a relatively high level. *(2) Discrimination*: this means that the similarity of the hashes of a pair of images with different content should be very low. *(3) Key-dependent security:* image hashing should be key-dependent in order to enhance the security of the algorithm. In the case where a key is unknown, an image hash produced with the wrong key should be entirely different from the original, meaning that an attacker cannot crack or forge the image hash without the secret key. In general, a typical image hashing scheme can be divided into three main stages: pre-processing, feature extraction, and hash representation, as shown in Fig 1.

Although many existing methods have been designed for grayscale images, the content conveyed by color information is also very important. In this paper, both the textural features and the color features of an image are considered, and these are combined to generate a synthetic image hashing sequence that can achieve a good balance between robustness and discrimination. The contributions of our scheme can be briefly summarized as the following two aspects:

(1) Texture features: After applying a discrete wavelet transform (DWT) to a grayscale image, we use low-frequency components to extract textural information, which can reduce the influence of noise and blur. Compared with the traditional local binary pattern (LBP) method, a texture image obtained using a Weber local binary pattern (WLBP) is less affected by noise. In addition, we use a global feature histogram to represent textural information and further reduce the data.

(2) Color features: The color angle is extracted using two color angular patterns. The color angle values from adjacent pixels have good correlation and redundancy, and this approach can enhance the robustness after dimension reduction. In addition, the energy concentration aspect of the discrete cosine transform (DCT) coefficient is utilized, leading to a high level of robustness.

After the extraction of the texture and color features, discrimination can be improved when cascading the hash of these two features. The experimental results show that image content modifications such as image addition, deletion, color transformation or copy-paste can be successfully authenticated with the proposed scheme.

The remaining parts of the paper are organized as follows. Section II reviews the state-of-the-art works. Section III describes the proposed image hashing scheme detailedly, including the processes of pre-processing, perceptual feature extraction, and hash generation. Experimental results and comparisons are presented in Section IV. Finally, Section V draws a conclusion of this paper.

II. RELATED WORKS

At present, there have been many ways being proposed to generate image hashing, which is represented by binary or decimal format. Almost all image hashing methods can be generally classified into two categories: feature extraction and dimensionality reduction.

A. FEATURE-EXTRACTION-BASED HASHING ALGORITHMS

In terms of feature extraction, many early methods developed image hashing algorithms based on the use of frequency domain transformations such as the DCT and discrete Fourier transform (DFT). Fridrich and Goljan [5] used a DCT to project an image onto random smooth patterns, and then used a threshold to quantize each pattern to obtain several bits in the form of image hashing. Although this scheme is robust to color adjustment and filtering, it fails to resist the modifications of rotation, shift and scale. Tang *et al*. [6] extracted color features through color vector angle (CVA). The inscribed circle of the normalized color image is selected to generate histograms. Finally, DCT, DWT and DFT are applied to obtained shorter hash sequences respectively. The results show the DCT scheme is better than the other two.

In addition to conventional frequency features, numerous other transforms have been applied to construct image hashes. Ouyang *et al*. [7] converted the rotation transform into a translation transform by using a log polar transform. The final image hashing was generated according to the correlation of the low-frequency magnitude coefficients of a quaternion discrete Fourier transform (QDFT), which were extracted from images in advance. Tang *et al*. [8] proposed the amplitude of the 1-D DFT coefficient is invariant to translation after log polar transformation, which can achieve the purpose of rotation invariance. After dimensionality reduction by the multidimensional scaling method, the similarity between the features of the secondary image is retained. However, this

scheme is sensitive to noise due to the severe image deformation after the log polar transformation. Li *et al*. [9] proposed a robust image method to resist rotation and quantification by using a Gabor filter and a dithered lattice vector quantity (LVQ). Swaminathan *et al*. [10] proposed a method for extracting a hash by using a Fourier-Mellin transform, which is resilient to moderate geometric, JPEG compression and filtering distortions. Lefebvre *et al*. [11] proposed a hashing algorithm that used radon transforms to maintain robustness to geometric transformation. Jin *et al*. [12] proposed a hashing scheme based on a Radon transform that can better resist affine transformation. Wu *et al*. [13] proposed a print/scan-resistant image hashing algorithm based on the Radon and wavelet transforms. Li *et al*. [14] proposed a hashing algorithm based on quaternion polar cosine transform (QPCT), which can holistically capture complementary visual features.

Spatial domain features have also been extracted directly by other methods. Schneider *et al*. [15] designed an image hashing scheme that used histogram features. Monga *et al*. [16] proposed the use of image feature points in image hashing, in which the geometry-preserving feature points were extracted by an iterative feature detector. In another work, Yang *et al*. [17] proposed a method for video copy retrieval by applying speeded up robust features (SURF). The local features are extracted using SURF in each divided block, in a frame-by-frame manner. The hashing was then generated by the difference between the SURF points of adjacent blocks in the Hilbert curve. Qin *et al*. [18] utilized textural features based on the dual-cross pattern (DCP) and salient structural features extracted from the frequency coefficients and position information to form a hash value. In order to improve robustness, Tang *et al*. proposed a ring partition method to resist rotation [19]. The preprocessed image was divided into a number of rings of the same area before calculating the statistical characteristics of the mean, variance, skewness and kurtosis of each ring. The vector distance was then selected as the hash value after the characteristics were compressed. Tang *et al*. [20] used the L* component in CIE L*a*b* color space to construct a three-dimensional tensor with the way of block partition. Then they obtained a valid low rank matrix through Tucker decomposition and generated the hash sequence by further compression and quantization. The defect of this method is that it cannot resist the rotation attack.

B. DIMENSIONALITY-REDUCTION-BASED HASHING ALGORITHMS

Other schemes are based on data dimensionality reduction. The classical dimensional reduction methods include singular value decomposition (SVD), principal component analysis (PCA), non-negative matrix factor (NMF), and locally linear embedding (LLE). A SVD-SVD algorithm was proposed by Kozat *et al*. [21] in which a secondary image was constructed by capturing semi-global geometric characteristics. The hashing is robust to severe perturbations, such as 50% cropping by area with 20° rotations, thereby achieving better classification. Ghouti [22] constructed image hashing by applying quaternion singular value decomposition (QSVD), in which the color correlation and interaction were taken into account. Based on the above motivation, Monga *et al*. [23] designed an NMF-NMF method to generate image hashing. By using this low-rank matrix factorization twice, the probability of error was minimized and misclassification was significantly reduced. Qin *et al*. [24] exploited SVD to construct a secondary image and conducted block truncation coding (BTC) to extract perceptual image features. In [24], a center-symmetrical local binary pattern (CSLBP) was computed and compressed to produce binary image hashing. Tang *et al*. [25] adopted the same ring partition method as in [19], and constructed a rotation-invariant matrix based on these rings. NMF was then applied to obtain a compact image hashing. Li *et al*. [26] presented a dictionary-learning algorithm to train a distortion-resistant sparse coding model for image hashing.

In this paper, we combine local texture and global color features to generate image hashing. Textural features applied with DWT and color features exploited with DCT are extracted to achieve better robustness and discrimination. In order to obtain shorter sequences, we use PCA to further reduce the amount of data, thereby generating a shorter image hash.

III. PROPOSED SCHEME

The proposed image hashing scheme consists of three main procedures: pre-processing, extraction of local texture and global color features, and hash generation. A flowchart depicting the main steps in our scheme is shown in Fig 2.

A. PRE-PROCESSING

In order to guarantee that the final hashes for different images have a fixed length, the size of input original image $\mathbf{I}_o$ is first normalized to $L \times L$ using a bilinear resampling operator, and the resized image is then processed using Gaussian low-pass filtering to alleviate the influences of noise and blurring. Gaussian filtering can be realized using a two-dimensional Gaussian function to form a convolution mask G:

$$G(u,v) = \frac{g(u,v)}{\sum_u \sum_v g(u,v)} = \frac{1}{2\pi\sigma^2} e^{-\frac{(u-K-1)^2+(v-K-1)^2}{2\sigma^2}} \quad (1)$$

where $G(u, v)$ denotes the value at coordinates (u, v) of the convolution mask; σ is the standard deviation of the Gaussian function g with a mask size of $(2K + 1)\times(2K + 1)$; and $g(u, v) = e^{-\frac{u^2+v^2}{2\sigma^2}}$. The normalized and filtered image, denoted as $\mathbf{I}$, can be obtained by applying the convolution mask G to the normalized image. Thus, $\mathbf{I}$ can be seen as the result after pre-processing, which then undergoes perceptual feature extraction.

B. PERCEPTUAL FEATURE EXTRACTION

In an image hashing scheme, the produced hash must be closely related to the perceptual image features, which should

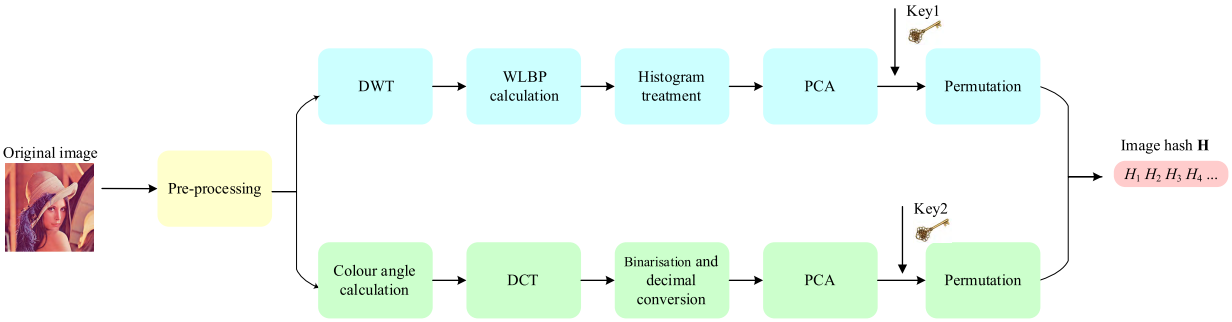

FIGURE 2. Flowchart of the proposed scheme.

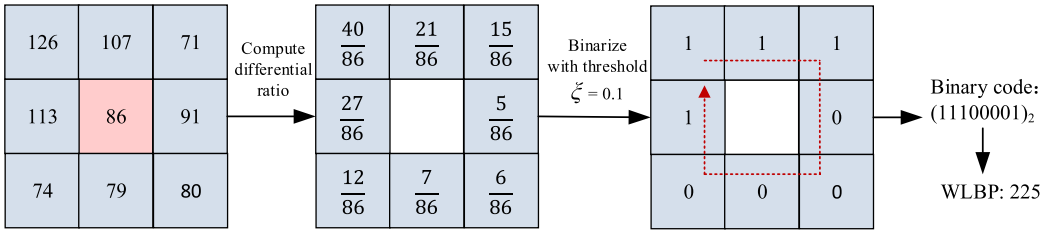

FIGURE 3. Example of WLBP calculation.

be able to reflect the principal contents and details of the input image. Since information about the texture and color is very important in perceptual descriptions of images, local binary pattern (LBP)-based local texture features and color angular patterns are involved in the feature extraction stage for our scheme, and these are discussed in the following.

1) LOCAL TEXTURE FEATURE EXTRACTION BASED ON WEBER'S LAW

The LBP is a simple and effective descriptor to represent a local texture feature [27], and often represents the pixels of an image as a series of binary numbers by comparing the central pixel with its eight neighbors in the window used. In the traditional LBP descriptor, it is easy to calculate the sign and magnitude information of local differences, as shown in (2) and (3).

$$LBP_{P,R}(x_c) = \sum_{n=0}^{P-1} S(x_n - x_c) \times 2^n \tag{2}$$

$$S(x) = \begin{cases} 1, x \geq 0 \\ 0, x < 0 \end{cases} \tag{3}$$

where x_c is the gray value of a given center pixel, and x_n ($n = 0, 1, \ldots, P-1$) represent the gray values of the P neighbors, distributed in a circular window with radius R for x_c.

However, the traditional LBP descriptor does not consider the difference between the neighborhood pixels and the central pixel, and may therefore even generate the same binary code for visually different contents. In addition, since it is sensitive to noise, small changes to the central or neighborhoods pixels will lead to a drastically different LBP value. In order to represent local textural features more accurately, we integrate human visual characteristics into the LBP descriptor in this work. We introduce Weber's law [28] to enhance the traditional LBP descriptor. Weber's law can be expressed as:

$$K = \frac{\Delta I}{I_{BG}} \tag{4}$$

where I_{BG} represents the initial stimulus; ΔI is the smallest increment or decrement required for the change to be perceived; and K is a constant defined as the Weber fraction. Equation (4) represents the relationship between physical quantities and psychological magnitudes. When I_{BG} is very large, a change smaller than ΔI in the stimulus cannot be perceived. However, under the relatively low stimulus of the environment, even a tiny change is enough to attract attention. Hence, the right-hand side of the equation must be greater than K to trigger human perception of changes. We apply Weber's law to a visual stimulus arising from image features to form the WLBP descriptor [29], which uses the difference ratio between the central pixel and surrounding pixels, as shown in (5) and (6).

$$WLBP_{P,R,\xi}(x_c) = \sum_{n=0}^{P-1} S_\xi \left(\frac{|x_n - x_c|}{x_c} \right) \times 2^n \tag{5}$$

$$S_\xi(x) = \begin{cases} 1, x \geq \xi \\ 0, x < \xi \end{cases} \tag{6}$$

where x_c, x_n, P and R have the same meanings as in (2), and ξ is a threshold for the difference in gray level. For a more intuitive comprehension of WLBP, Fig 3 illustrates an example of the calculation of WLBP, where $x_c = 86$, $P = 8$, $R = 1$, and $\xi = 0.1$.

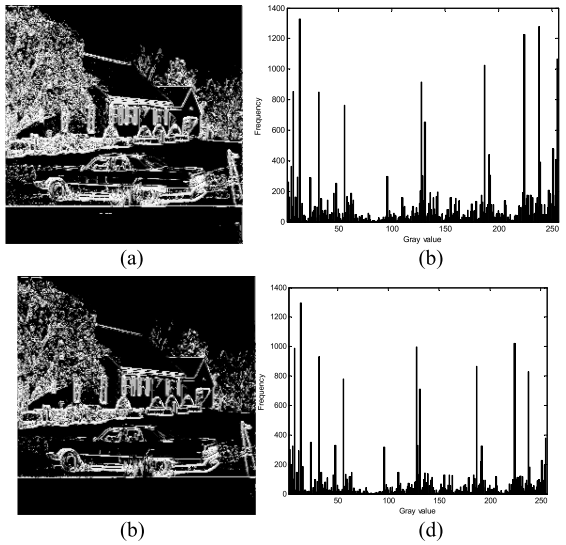

FIGURE 4. WLBP maps and corresponding histograms with different thresholds ξ for House. (a) WLBP map with $\xi = 0.1$; (b) WLBP map with $\xi = 0.15$; (c) histogram of (a); (d) histogram of (b).

In order to effectively apply a WLBP descriptor to extract local texture features for our image hashing, the following steps are carried out. First, the normalized, filtered color image $\mathbf{I}$ with size $L \times L$ is converted into a gray image $\mathbf{I_g}$ for WLBP extraction. Following this, DWT is performed on $\mathbf{I_g}$ (here, the wavelet base db2 is used), and the low frequency sub-band LL_1 is then utilized to calculate WLBP, using (5). As a result, we can obtain a WLBP map $\mathbf{W}$ of size $M \times M$ that consists of decimal values, where $M = (L/2 - 2)$. It is worth noting that we do not consider the four boundary pixels of LL_1 as central pixels when applying WLBP. Then, the histogram of $\mathbf{W}$, denoted by $\mathbf{U}$, can be derived as

$$\mathbf{U}(k) = \sum_{x=1}^{M}\sum_{y=1}^{M} Z(WLBP_{P,R,\xi}(x_c), k), k \in \{1, 2, \ldots, 255\} \tag{7}$$

$$Z(u, v) = \begin{cases} 1, & \text{if } u = v \\ 0, & \text{otherwise} \end{cases} \tag{8}$$

Note that, the value of $\mathbf{U}(k)$ is considerably higher than the other values in the case where $k = 0$, and the value of $\mathbf{U}(0)$ does not have a great deal of representational significance. Hence, in our algorithm, we do not take $\mathbf{U}(0)$ into consideration in the final histogram $\mathbf{U}$. Fig 4 illustrates examples of WLBP maps and their corresponding histograms for the image *House*, with different values of the threshold ξ, where the parameters P and R are set to 8 and 1, respectively. Fig 4 (a) and (b) show the WLBP maps with $\xi = 0.1$ and $\xi = 0.15$, respectively, and (c) and (d) are their corresponding histograms. It can be observed from Fig 4 that the texture has a different level of sharpness for different thresholds. When $\xi = 0.1$, the textural structure is clearer and the distribution of histograms is flatter; for $\xi = 0.15$, the textural details are reduced in the map, leading to lower histogram values.

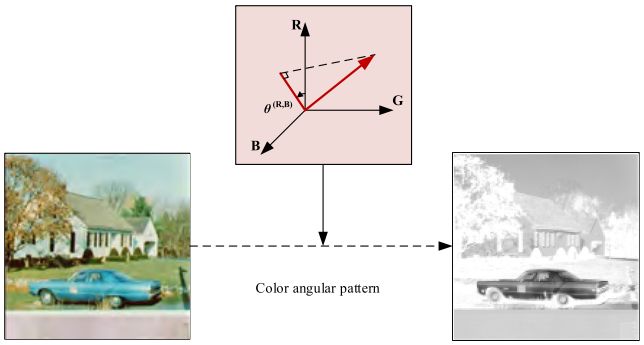

FIGURE 5. Schematic diagram of color angular matrix using a color angular pattern from the two color components of R and B.

After obtaining the histogram $\mathbf{U}$ for the WLBP map $\mathbf{W}$, the PCA dimension reduction method is applied to $\mathbf{U}$ to produce $\mathbf{U}$' with length L_1. We then normalize each value in $\mathbf{U}$' to the range 0–255 in order to guarantee the uniformity of the final hash value generated with different features. In order to enhance security, we permute half of $\mathbf{U}$' at random to get a matrix $\mathbf{X}$, using the secret key K_1. Note that the permutation of this permutation is $A_{L_1}^{L_1/2}$. If $L_1 = 64$, the number of possible orderings is $A_{64}^{32} = 4.82 \times 10^{53}$.

2) COLOR ANGLE FEATURE

Different colors may provide a different understanding of the image contents. Hence, the color features of images are also considered in our hashing scheme. Specifically, in our scheme, two components from the three RGB channels in the color image $\mathbf{I}$ are selected as in [30] to calculate the color angle $\theta_{i,j}^{(p,q)}$, as shown in (9):

$$\theta_{i,j}^{(p,q)} = \tan^{-1}\left(\frac{p_{i,j}}{q_{i,j} + \varepsilon}\right) \tag{9}$$

where p and q denote two selected components from RGB three channels, respectively, and ε is a constant with a small value to avoid a singularity when the denominator in (9) is zero. Therefore, for each pixel at (i, j), $\theta_{i,j}^{(p,q)}$ is close to the color angle of its neighboring points, which is different from that of the pixel at (i, j) in a different image. Fig 5 presents a schematic diagram of the calculation of the color angular pattern, where the color components of the R and B channels are used.

When the color angles for all pixels have been calculated, a matrix $\mathbf{A}$ consisting of $N \times N$ values can be obtained. A DCT is then applied to the color angle matrix $\mathbf{A}$, and for the first N_1 coefficients in a zigzag scanning order, a banalization process is implemented to obtain a binary sequence $\mathbf{V}$ with length N_1, as shown in (10).

$$\mathbf{V}(i) = \begin{cases} 1 & \text{if } c(i) \geq \mu \\ 0 & \text{otherwise} \end{cases} \quad (1 \leq i \leq N_1) \tag{10}$$

where $c(i)$ denotes the i^{th} coefficient in a zigzag scanning order, and μ denotes the average value of all N^2 DCT coefficients for $\mathbf{A}$. Then, every eight binary bits in $\mathbf{V}$ are converted to a decimal number to adjust the value of this number being

within the range of [0, 255]. In order to reduce the dimensions, PCA is also performed on these $N_1/8$ decimal values to obtain $\mathbf{V}$' with length L_2. We encrypt $\mathbf{V}$' with a secret key K_2 in the same way as K_1. We obtain a matrix $\mathbf{Y}$ by randomizing half of $\mathbf{V}$' to achieve encryption. The permutation of this encryption is $A_{L_2}^{L_2/2}$, and for L_2 =26, the number of possible orderings is $A_{26}^{13} = 6.48 \times 10^{16}$.

C. HASH GENERATION

After acquiring the two feature vectors, i.e., $\mathbf{X}$ of WLBP-based texture feature and $\mathbf{Y}$ of color angle based feature, we concatenate $\mathbf{X}$ and $\mathbf{Y}$ to generate the sequence $\mathbf{H}$ consisting of $L = L_1 + L_2$ decimal digits. $\mathbf{H}$ is referred to as the final image hashing of original image $\mathbf{I}_o$.

IV. EXPERIMENTAL RESULTS AND COMPARISONS

In order to evaluate the effectiveness and superiority of the proposed scheme, a large number of experiments and comparisons were conducted with respect to perceptual robustness, discrimination and key-dependent security. All experiments were implemented on a computer with a 3.20 GHz Core i5-6500 CPU, 4.00 GB memory, and Windows 7 operating system, and the programming environment was MATLAB 2014a.

A. SIMILARITY MEASUREMENT AND PARAMETER SETTING

In our experiments, the similarity between two image hashes, $\mathbf{H}^{(1)}$ and $\mathbf{H}^{(2)}$, was measured by Euclidean distance:

$$D = \frac{1}{D_{\max}} \sqrt{\sum_{i=1}^{L} \left| H_i^{(1)} - H_i^{(2)} \right|^2} \tag{11}$$

where the function D denotes the normalized Euclidean distance between two image hashes $\mathbf{H}^{(1)} = [H_1^{(1)}, H_2^{(1)}, \ldots, H_L^{(1)}]$ and $\mathbf{H}^{(2)} = [H_1^{(2)}, H_2^{(2)}, \ldots, H_L^{(2)}]$, D_{max} is the maximum Euclidean distance obtained in the experiment. If the distance between two hashes is smaller than a pre-determined threshold T, the two corresponding images can be judged as visually similar, otherwise as visually distinct.

There are several parameters in the three main stages of the proposed image hashing scheme, and we explicitly describe parameter setting used in our experiments as follows. During the stage of pre-processing, the size of input images was normalized to 256 × 256, i.e., $I_1 = 256$, and the standard deviation σ of Gaussian low-pass filtering was set to 1. During the stage of perceptual feature extraction, as for local texture feature, the threshold ξ of WLBP calculation was set to 0.1, the size of WLBP map is 126 × 126, i.e., $M = 126$, and the length L_1 of WLBP-based texture feature vector $\mathbf{X}$ was set to 64 digits; as for color angle feature, the number N_1 of selected DCT coefficients for color angle matrix $\mathbf{A}$ sized 256 × 256 in zig-zag scanning order was 16384, and the length L_2 of color angle based feature $\mathbf{Y}$ was set to 26 digits. As a result, the final hash $\mathbf{H}$ consists of 90 digits, i.e., $L = L_1 + L_2 = 90$. A summary of parameter setting is listed in Table 1.

(a) (b) (c) (d) (e)

FIGURE 6. Standard test images. (a) ***Airplane***, (b) ***Baboon***, (c) ***House***, (d) ***Lena***, and (e) ***Peppers***.

TABLE 1. Parameter Settings of The Proposed Scheme.

Stages	Parameters	Values
Pre-processing	Image size $I_1 \times I_1$	256 × 256
	Standard deviation σ	1
WLBP texture feature	Threshold ξ	0.1
	WLBP map size $M \times M$	126 × 126
	Length L_1	64 digits
Color angle feature	Coefficient number N_1	16384
	Length L_2	26 digits
Hash generation	Final hash length L	90 digits

B. PERCEPTUAL ROBUSTNESS

Fig 6 shows five standard color images sized 256 × 256, i.e., *Airplane*, *Baboon*, *House*, *Lena*, and *Peppers*. All five images can be downloaded from [31]. Eight commonly content-preserving operations, including brightness adjustment, contrast adjustment, Gaussian low-pass filtering, JPEG compression, Gamma correction, scaling, salt and pepper noise, and speckle noise were conducted to produce visually similar versions with respect to original images in Fig 6 for robustness evaluation. Detailed information of these eight content-preserving operations with corresponding parameter setting are listed in Table 2.

For each image in Fig 6, we first generated the hashes for itself and its visually similar versions produced by each of content-preserving operations in Table 2, and then calculated the normalized Euclidean distances according to (11). Thus, 64 normalized Euclidean distances D can be obtained between the hash pairs of each original test image and its 64 visually similar versions, and for the five original images in Fig 6, totally 320 normalized Euclidean distances can be produced. In each subfigure of Fig 7, the abscissa is parameter value of each kind of content-preserving operations, and the ordinate is the normalized Euclidean distance between the hash pair of original image and corresponding similar version. It can be observed from Fig 7 that, the values of normalized Euclidean distances are generally smaller than 0.25. That is to say, our image hashing scheme achieves a satisfactory performance of perceptual robustness against commonly-used content-preserving operations.

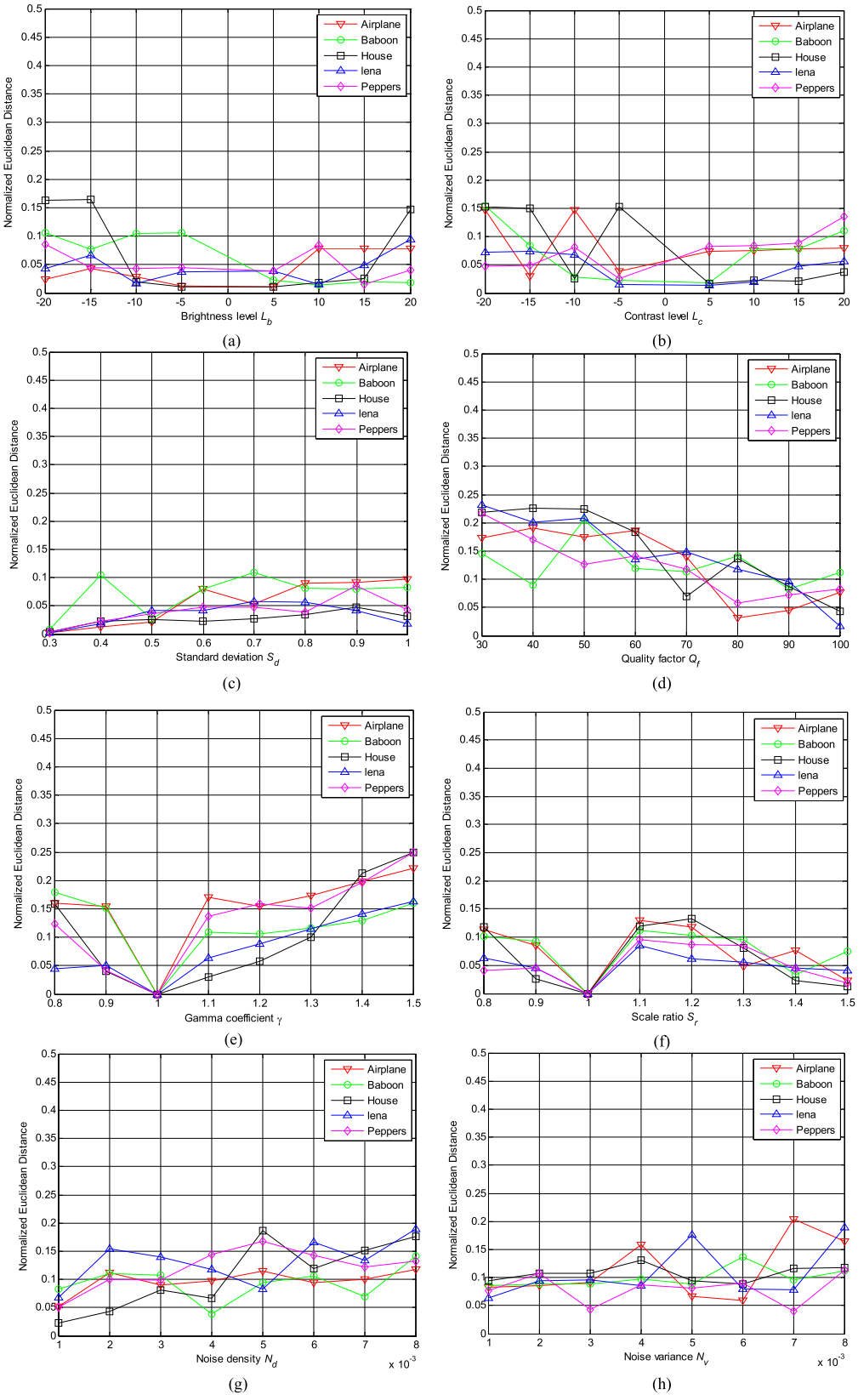

FIGURE 7. Performance of perceptual robustness on five standard test images. (a) Brightness adjustment. (b) Contrast adjustment. (c) Gaussian low-pass filtering. (d) JPEG compression. (e) Gamma correction. (f) Image scaling. (g) Salt and pepper noise. (h) Specklenoise.

TABLE 2. Eight content-preserving operations with corresponding parameter setting.

Operations	Parameter descriptions	Parameter values	Number of generated images
Brightness adjustment	Brightness level L_b	±5, ±10, ±15, ±20	8
Contrast adjustment	Contrast level L_c	±5, ±10, ±15, ±20	8
Gaussian low-pass filtering	Standard deviation S_d	0.3, 0.4, …, 1.0	8
JPEG compression	Quality factor Q_f	30, 40, …, 100	8
Gamma correction	Gamma coefficient γ	0.8, 0.9, …, 1.5	8
Image scaling	Scale ratio S_r	0.8, 0.9, …, 1.5	8
Salt and pepper noise	Noise density N_d	0.001, 0.002, …, 0.008	8
Speckle noise	Noise variance N_v	0.001, 0.002, …, 0.008	8

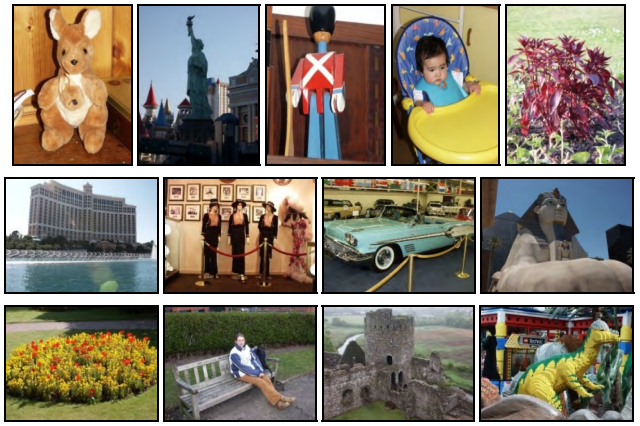

FIGURE 8. Some representative images in UCID.

C. DISCRIMINATION CAPABILITY

In order to evaluate the discrimination capability (i.e., anti-collision) of our hashing scheme, the uncompressed color image database (UCID) [32] including 1338 various images with the sizes of 512 × 384 and 384 × 512 was adopted in the experiments. Several representative images in UCID are presented in Fig 8.

We first generated hashes for all 1,338 color images in UCID, and then, for each image, the normalized Euclidian distances D was calculated between its hash and the other 1,337 hashes. As a result, there are totally (1338 × 1337)/2 = 894,453 normalized Euclidian distances D that can be obtained for all hash pairs from these 1,338 visually different images in UCID. Fig 9 shows the histogram of these 894,453 normalized Euclidian distances D, in which the abscissa denotes the value of normalized Euclidian distance D and the ordinate denotes corresponding occurrence frequency of D. It can be found that, the distribution of histogram for normalized Euclidian distances approximately conforms to a normal distribution with mean value $\mu = 0.6481$ and standard deviation $\sigma = 0.0782$ based on the parameter estimation.

Statistically speaking, the collision probability P_c for two perceptual different images is the probability that their normalized Euclidian distance D is smaller than a pre-determined threshold T, as shown in (12).

$$P_c(T) = \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{T}\exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right]dx = \frac{1}{2}\mathrm{erfc}\left(-\frac{T-\mu}{\sqrt{2}\sigma}\right) \tag{12}$$

TABLE 3. Collision Probabilities under Different Thresholds T.

Threshold T	Collision probability P_c
0.20	5.0917×10^{-9}
0.22	2.2250×10^{-8}
0.24	9.1242×10^{-8}
0.25	1.8043×10^{-7}
0.26	3.5120×10^{-7}
0.28	1.2690×10^{-6}
0.30	4.3055×10^{-6}
0.32	1.3719×10^{-5}

where exp(·) represents the exponential function with the natural constant e as its base and erfc(·) denotes the complementary error function. Thus, we can calculate the collision probabilities P_c with different pre-determined thresholds T through (12), as listed in Table 3. Obviously, the smaller the threshold T is set, the smaller the collision probability P_c is. On the other hand, since the normalized Euclidian distance D between the hash pair of two visually similar images should be smaller than the threshold T, hence, smaller threshold T may influence the performance of perceptual robustness. It can be observed from Fig 7 that, the normalized Euclidian distances D of our scheme under the eight kinds of commonly-used content-preserving operations are almost all smaller than 0.25. In addition, when T is equal to 0.25, the collision probability of our image hashing scheme is 1.8043×10^{-7}, which is small enough for the application of retrieval and authentication for images. Therefore, in the proposed scheme, we can set the threshold T as 0.25 to achieve a satisfactory compromise for the two performances of perceptual robustness and discrimination simultaneously.

D. KEY-DEPENDENT SECURITY

As stated in Section I, an image hashing scheme should satisfy that different secret keys produce significantly distinct image hashes. In our hashing scheme, two secret keys K_1 and K_2 are applied in the final stage, i.e., hash generation. Specifically, K_1 has A_{64}^{32} possibilities and K_2 has A_{26}^{13} possibilities. The number of possible concatenated hashing is equivalent to the multiplication $A_{64}^{32} \times A_{26}^{13}$, so the adversaries are very difficult to crack the hash without knowing the correct keys. The performance of key-dependent security of our hashing scheme is demonstrated in Fig 10, where the abscissa is the index of the 100 groups of randomly-generated wrong

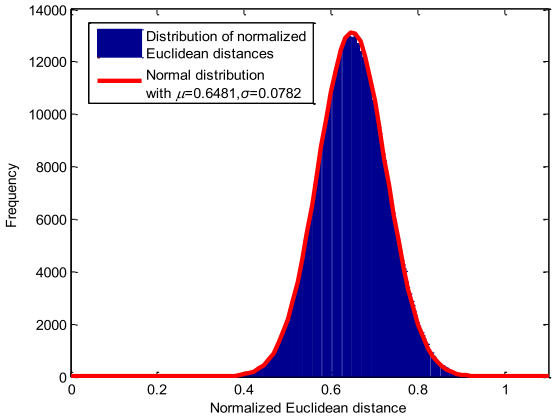

FIGURE 9. Distribution of 894,453 normalized Euclidean distances between hash pairs of the 1338 visually distinct images in UCID.

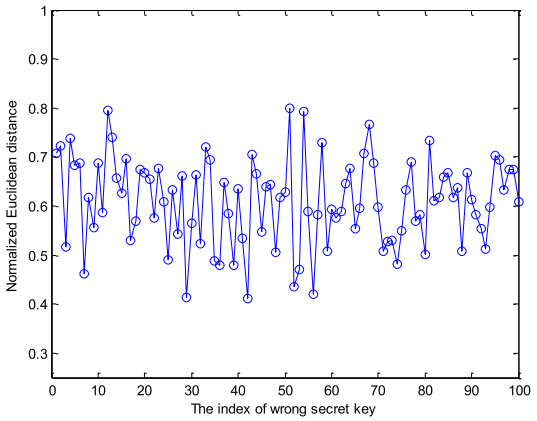

FIGURE 10. Normalized Euclidean distances between hash pairs derived from the correct secret key and 100 wrong secret keys.

secret keys K_1 and K_2, and the ordinate is the corresponding normalized Euclidian distance between the hash pairs derived from the correct and the wrong secret keys for Baboon. It can be observed that, almost all normalized Euclidian distances in Fig 10 are located in the range of [0.4, 0.8], and it means that it is significantly difficult for the adversaries to regenerate the same hash without knowing the correct keys. Hence, the security of the proposed scheme completely depends on the secret keys and satisfies the security requirement in the sense of cryptography.

E. COMPARISONS WITH STATE-OF-THE-ART SCHEMES

In order to demonstrate the superiority of the proposed scheme, we compare our hashing scheme with five state-of-the-art schemes, i.e., RP-IVD hashing [19], BTC-CSLBP [24], RP-NMF hashing [25], CV-ED hashing [33], and CVA-LLE hashing [34]. Since the discrimination capability is often contradictory with perceptual robustness for image hashing schemes, thereby, in order to fairly conduct the comparisons for the proposed scheme and [19], [24], [25], [33], [34], we considered the comprehensive performances of perceptual robustness and discrimination as the classification capability of differentiating visually similar and distinct images.

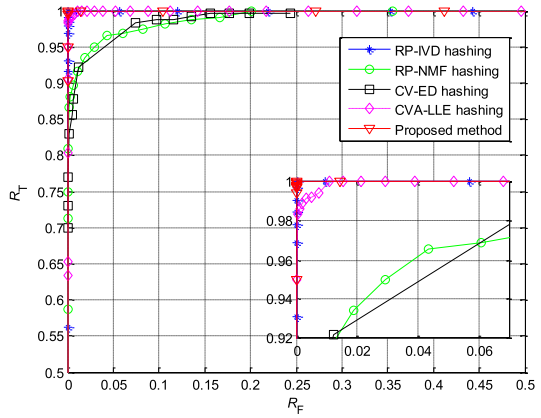

FIGURE 11. ROC curves of classification capability for the proposed scheme, [19], [25], [33], and [34].

1) COMPARISON OF RECEIVER OPERATING CHARACTERISTIC CURVE FOR CLASSIFICATION

Two typical indices, i.e., the true positive rate R_T and the false positive rate R_F, were utilized for performance evaluation:

$$R_T = \frac{N_{true}}{N_{same}} \tag{13}$$

$$R_F = \frac{N_{false}}{N_{different}} \tag{14}$$

where N_{true} denotes the number of actually similar images correctly classified as similar images, N_{false} denotes the number of actually distinct images falsely classified as similar images, and N_{same} and $N_{different}$ are the total number of actually similar and distinct images, respectively. Fig 11 shows the receiver operating characteristic (ROC) curves of our scheme, [19], [25], [33], and [34] to demonstrate comparison results for classification capability based on overall performances of perceptual robustness and discrimination, in which the abscissa and ordinate denote true positive rate P_T and false positive rate P_F, respectively, and there are 894,453 visually distinct image hash pairs and similar image pairs used as our experimental data. Obviously, the closer to top-left corner of Fig 11 the ROC curve is, the better classification capability the image hashing scheme achieves. It can be observed from Fig 11 that, ROC curve of the proposed scheme is much closer to the top-left corner than those of the schemes [25], [33], [34]. That is to say, the proposed image hashing scheme has better capability of image classification compared with the reported schemes [25], [33], [34].

2) COMPARISON OF INTEGRATED HISTOGRAM

Besides the ROC curves in Fig 11, the integrated histograms of normalized Euclidean distances for visually similar and distinct images were also utilized for performance comparison. The subfigures (a), (b), (c) and (d) in Fig 12 illustrate the integrated histograms corresponding to the schemes, [25], [34], [24] and our scheme, and the blue and red stems in each subfigure represent the distributions of normalized Euclidean distances for visually similar images

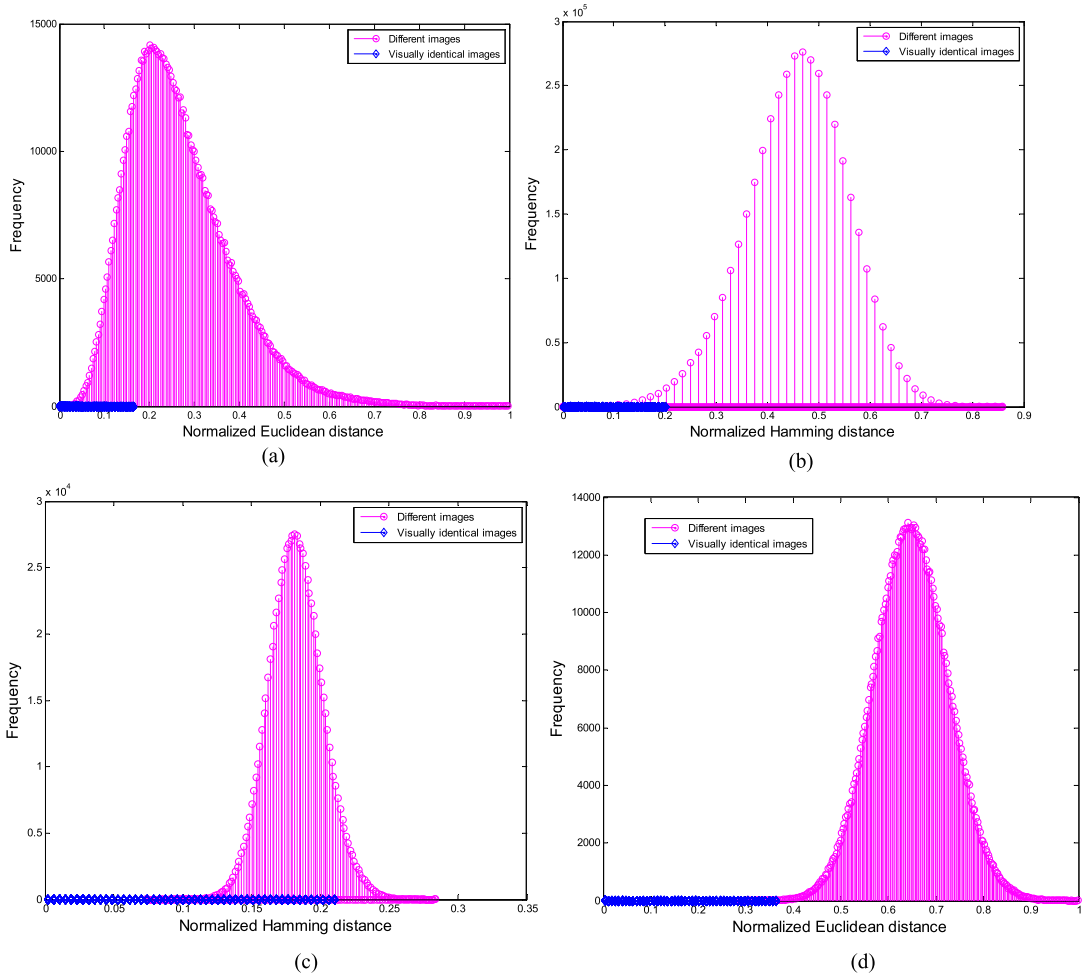

FIGURE 12. Classification performance comparisons using integrated histograms of normalized distances. The blue and the red stems correspond to the distributions of normalized distances for visually similar images and visually distinct images, respectively. (a) Scheme in [25], (b) Scheme in [34], (c) Scheme in [24], (d) Proposed scheme.

TABLE 4. Comparisons of Execution Time and Hash Length for Our Scheme and [19], [24], [25], [33], and [34].

Complexity Performance	RP-NMF	RP-IVD	CV-ED	CVA-LLE	BTC-CSLBP	Our scheme
Execution time	12.8 s	0.61 s	14.6 s	2.0 s	1.1 s	2.5 s
Hash length	64 digits	40 digits	40 digits	64 bits	672 bits	90 digits

(generated from content-preserving operations in Table 2 for the images shown in Fig 6) and visually distinct images (from UCID database), respectively. As observed in each histogram, the blue and red stems distribute in the left and right sides of the horizontal axis. Considering from the classification performance, the less overlap, the better the scheme. As shown in Fig 12, the overlapping intervals of the comparison schemes [25], [34], [24] and our scheme are [0.0138, 0.1653], [0, 0.2031], [0.0729, 0.2128], and [0.2867, 0.3676] respectively. The proportions of overlapping parts for [25], [34], [24] and our scheme are 0.1515, 0.2031, 0.1399, and 0.0809, respectively, which also demonstrates the superiority of the classification performance of our scheme.

3) COMPARISON OF COMPUTATIONAL COMPLEXITY

We also conducted complexity comparisons with respect to execution time and hash length for our hashing scheme and the schemes [19], [24], [25], [33], [34]. The execution time of producing image hashes for 100 images in UCID were recorded, and then the average time of hash calculation for an image can be obtained, as listed in Table 4. We can find from Table 4 that, the average execution time of the proposed scheme is faster than those of the schemes in [25], [33], but is slightly slower than those of the schemes in [19], [24], [34]. In addition, with respect to hash length, our scheme is longer than the other schemes.

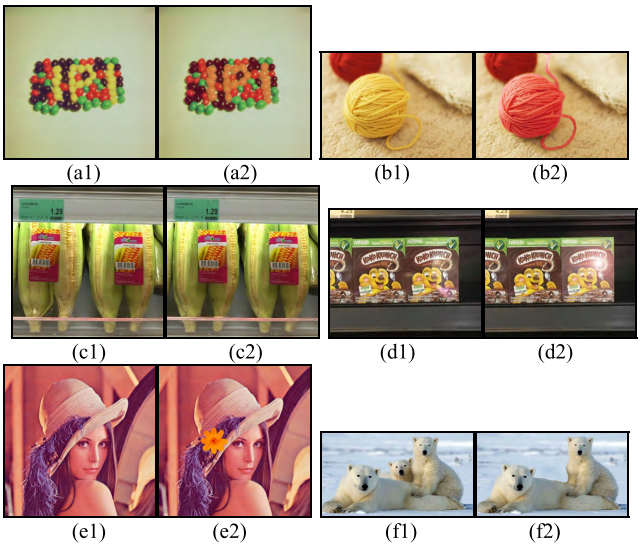

FIGURE 13. Original images and corresponding tampered versions. (a1-f1) are original images; (a2) and (b2) are local color modified versions of (a1) and (b1), respectively; (c2) is the local copy and paste version of (c1); (d2) is the local copy and paste, and lens flare addition version of (d1); (e2) is the external object insertion version of (e1); and (f2) is the internal object remove version of (f1).

TABLE 5. Normalized Euclidean Distances between Original Images and Their Tampered Versions.

Images	Normalized Euclidean Distances
(a1) and (a2)	0.3375
(b1) and (b2)	0.3714
(c1) and (c2)	0.3711
(d1) and (d2)	0.3428
(e1) and (e2)	0.3994
(f1) and (f2)	0.3370

Based on the above comparison results, it can be concluded that, the proposed scheme not only achieves a superior capability of image classification based on perceptual robustness and discrimination compared with some state-of-the-art schemes, but also guarantees a satisfactory security.

F. APPLICATION OF IMAGE AUTHENTICATION

Image hashing can be applied for image authentication by calculating the distance between the hashes of original image and the potential tampered image, and then comparing the distance with the pre-determined threshold to determine whether the image is tampered or not. Obviously, the higher the accuracy of tampering detection, the better the performance of the scheme. Typical tampering manipulations include object insertion and removal, and color modification, etc. In order to demonstrate the capability of image authentication of our scheme, we chose six images as shown in Fig 13(a1), (b1), (c1), (d1), (e1), and (f1), respectively, and conducted tampering manipulations on them with the aid of software *Adobe Photoshop*. Their corresponding tampered results are shown in Fig 13(a2), (b2), (c2), (d2), (e2), and (f2), respectively, in which (a2) and (b2) are local color change, (c2) is local copy and paste operation, (d2) is local copy and paste operation and lens flare addition, (e2) is external object insertion, and (f2) is internal object remove, respectively. Table 5 lists the normalized Euclidean distances between original images and their corresponding tampered versions. We can find that, all normalized Euclidean distances in Table 5 are greater than the threshold $T = 0.25$, which means the tampering manipulations caused significant changes on image hashes and our scheme can achieve satisfactory performance in the application of image authentication.

V. CONCLUSIONS

In this paper, we propose a robust image hashing scheme that incorporates the local texture and global color features of an image. For texture feature extraction, we use Weber's law, which is based on the difference ratio between the central pixels and their surrounding pixels; that is, the WLBP value of each low-frequency wavelet coefficient is calculated, before a histogram count and a PCA dimensionality reduction operation are carried out. For color feature extraction, the color angular pattern is applied to each pixel to measure the correlations between the different color channels. The values of the color angles are then dimensionally reduced using a DCT and significant coefficients are selected before the two features are cascaded to generate an integrated hash. Experiments were conducted to verify the efficacy of the proposed scheme, and to demonstrate the robustness of our hashing method against many common content-preserving operations such as global brightness adjustment, contrast adjustment, Gaussian low-pass filtering, JPEG compression, Gamma correction, image scaling, salt-and-pepper noise, and speckle noise. Comparison with certain state-of-the-art schemes, the ROC curve and comprehensive classification performance demonstrate the superiority of the proposed hashing scheme in terms of its perceptual robustness and discriminative ability. In addition, the proposed scheme can also be applied to image content authentication.

However, there are still some potential avenues for improvement in our scheme, such as its robustness against small angle rotation, forensic capabilities for forgery detection with small region tampering, and improving its performance in terms of computation efficiency, and we will focus on these in future work.

ACKNOWLEDGMENT

The authors would like to thank the anonymous reviewers for their valuable suggestions.

REFERENCES

[1] C. D. Vleeschouwer, J.-F. Delaigle, and B. Macq, "Invisibility and application functionalities in perceptual watermarking an overview," *Proc. IEEE*, vol. 90, no. 1, pp. 64–77, Jan. 2002.
[2] C. Qin, P. Ji, X. Zhang, J. Dong, and J. Wang, "Fragile image watermarking with pixel-wise recovery based on overlapping embedding strategy," *Signal Process.*, vol. 138, pp. 280–293, Sep. 2017.
[3] C. Qin, P. Ji, C.-C. Chang, J. Dong, and X. Sun, "Non-uniform watermark sharing based on optimal iterative BTC for image tampering recovery," *IEEE MultiMedia*, vol. 25, no. 3, pp. 36–48, Jul./Sep. 2018.
[4] X. Zhu, X. Li, S. Zhang, Z. Xu, L. Yu, and C. Wang, "Graph PCA hashing for similarity search," *IEEE Trans. Multimedia*, vol. 19, no. 9, pp. 2033–2044, Sep. 2017.

[5] J. Fridrich and M. Goljan, "Robust hash functions for digital watermarking," in *Proc. Int. Conf. Inf. Technol., Coding Comput.*, Mar. 2000, pp. 178–183.
[6] Z. Tang, X. Li, X. Zhang, S. Zhang, and Y. Dai, "Image hashing with color vector angle," *Neurocomputing*, vol. 308, pp. 147–158, Sep. 2018.
[7] J. Ouyang, G. Coatrieux, and H. Shu, "Robust hashing for image authentication using quaternion discrete Fourier transform and log-polar transform," *Digit. Signal Process.*, vol. 41, pp. 98–109, Jun. 2015.
[8] Z. Tang, Z. Huang, X. Zhang, and H. Lao, "Robust image hashing with multidimensional scaling," *Signal Process.*, vol. 137, pp. 240–250, Aug. 2017.
[9] Y. Li, Z. Lu, C. Zhu, and X. Niu, "Robust image hashing based on random Gabor filtering and dithered lattice vector quantization," *IEEE Trans. Image Process.*, vol. 21, no. 4, pp. 1963–1980, Apr. 2012.
[10] A. Swaminathan, Y. Mao, and M. Wu, "Robust and secure image hashing," *IEEE Trans. Inf. Forensics Security*, vol. 1, no. 2, pp. 215–230, Jun. 2006.
[11] F. Lefebvre, J. Czyz, and B. Macq, "A robust soft hash algorithm for digital image signature," in *Proc. Int. Conf. Image Process.*, Sep. 2003, pp. II–495.
[12] J. S. Seo, J. Haitsma, T. Kalker, and C. D. Yoo, "A robust image fingerprinting system using the Radon transform," *Signal Process., Image Commun.*, vol. 19, no. 4, pp. 325–339, 2004.
[13] D. Wu, X. B. Zhou, and X. M. Niu, "A novel image hash algorithm resistant to print-scan," *Signal Process.*, vol. 89, no. 12, pp. 2415–2424, 2009.
[14] Y. N. Li, P. Wang, and Y. T. Su, "Robust image hashing based on selective quaternion invariance," *IEEE Signal Process. Lett.*, vol. 22, no. 12, pp. 2396–2400, Dec. 2015.
[15] M. Schneider and S.-F. Chang, "A robust content based digital signature for image authentication," in *Proc. 3rd IEEE Int. Conf. Image Process.*, Sep. 1996, pp. 227–230.
[16] V. Monga and B. L. Evans, "Robust perceptual image hashing using feature points," in *Proc. Int. Conf. Image Process.*, vol. 1, Oct. 2004, pp. 677–680.
[17] G. Yang, N. Chen, and Q. Jiang, "A robust hashing algorithm based on SURF for video copy detection," *Comput. Secur.*, vol. 31, no. 1, pp. 33–39, 2012.
[18] C. Qin, X. Q. Chen, X. Y. Luo, X. P. Zhang, and X. M. Sun, "Perceptual image hashing via dual-cross pattern encoding and salient structure detection," *Inf. Sci.*, vol. 423, pp. 284–302, Jan. 2018.
[19] Z. Tang, X. Zhang, X. Li, and S. Zhang, "Robust image hashing with ring partition and invariant vector distance," *IEEE Trans. Inf. Forensics Security*, vol. 11, no. 1, pp. 200–214, Jan. 2016.
[20] Z. Tang, L. Chen, X. Zhang, and S. Zhang, "Robust image hashing with tensor decomposition," *IEEE Trans. Knowl. Data Eng.*, vol. 31, no. 3, pp. 549–560, Mar. 2018.
[21] S. S. Kozat, R. Venkatesan, and M. K. Mihcak, "Robust perceptual image hashing via matrix invariants," in *Proc. Int. Conf. Image Process. (ICIP)*, vol. 5. Oct. 2004, pp. 3443–3446.
[22] L. Ghouti, "Robust perceptual color image hashing using quaternion singular value decomposition," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, May 2014, pp. 3794–3798.
[23] V. Monga and M. K. Mihçak, "Robust and secure image hashing via non-negative matrix factorizations," *IEEE Trans. Inf. Forensics Security*, vol. 2, no. 3, pp. 376–390, Sep. 2007.
[24] C. Qin, X. Chen, D. Ye, J. Wang, and X. Sun, "A novel image hashing scheme with perceptual robustness using block truncation coding," *Inf. Sci.*, vols. 361–362, pp. 84–99, Sep. 2016.
[25] Z. Tang, X. Zhang, and S. Zhang, "Robust perceptual image hashing based on ring partition and NMF," *IEEE Trans. Knowl. Data Eng.*, vol. 26, no. 3, pp. 711–724, Mar. 2014.
[26] Y. Li and L. Guo, "Robust image fingerprinting via distortion-resistant sparse coding," *IEEE Signal Process. Lett.*, vol. 25, no. 1, pp. 140–144, Jan. 2018.
[27] T. Ojala, M. Pietikäinen, and D. Harwood, "A comparative study of texture measures with classification based on featured distributions," *Pattern Recognit.*, vol. 29, no. 1, pp. 51–59, 1996.
[28] *Weber's Law of Just Noticeable Difference, University of South Dakota*. Accessed: Apr. 3, 2019. [Online]. Available: http://apps.usd.edu/coglab/WebersLaw.html
[29] Z. H. Xia, R. Lv, and X. M. Sun, "Rotation-invariant Weber pattern and Gabor feature for fingerprint liveness detection," *Multimedia Tools Appl.*, vol. 77, no. 14, pp. 18187–18200, 2018. doi: 10.1007/s11042-017-5517-9.
[30] S. H. Lee, J. Y. Choi, Y. M. Ro, and K. N. Plataniotis, "Local color vector binary patterns from multichannel face images for face recognition," *IEEE Trans. Image Process.*, vol. 21, no. 4, pp. 2347–2353, Apr. 2012.
[31] (2007). *USC-SIPI Image Database*. Accessed: Apr. 3, 2019. [Online]. Available: http://sipi.usc.edu/database/
[32] G. Schaefer and M. Stich, "UCID: An uncompressed color image database," *Proc. SPIE*, vol. 5307, pp. 472–480, Dec. 2004.
[33] Z. Tang, L. Huang, X. Zhang, and H. Lao, "Robust image hashing based on color vector angle and Canny operator," *AEU–Int. J. Electron. Commun.*, vol. 70, no. 6, pp. 833–841, 2016.
[34] Z. Tang, H. Lao, X. Zhang, and K. Liu, "Robust image hashing via DCT and LLE," *Comput. Secur.*, vol. 62, pp. 133–148, Sep. 2016.

CHUAN QIN received the B.S. degree in electronic engineering, and the M.S. degree in signal and information processing from the Hefei University of Technology, Anhui, China, in 2002 and 2005, respectively, and the Ph.D. degree in signal and information processing from Shanghai University, Shanghai, China, in 2008. Since 2008, he has been with the faculty of the School of Optical-Electrical and Computer Engineering, University of Shanghai for Science and Technology, where he is currently a Professor. He was with Feng Chia University, Taiwan, as a Postdoctoral Researcher, from 2010 to 2012. His research interests include image processing and multimedia security. He has published over 110 papers in these research areas.

YECEN HU received the B.S. degree in communication engineering from the University of Shanghai for Science and Technology, Shanghai, China, in 2017, where she is currently pursuing the M.S. degree in signal and information processing. Her research interests include image hashing, image retrieval, and image authentication.

HENG YAO (S'11–M'12) received the B.S. degree from the Hefei University of Technology, China, in 2004, the M.S. degree from Shanghai Normal University, China, in 2008, and the Ph.D. degree from Shanghai University, China, in 2012. He is currently with the School of Optical-Electrical and Computer Engineering, University of Shanghai for Science and Technology, China, where he is also an Associate Professor. His research interests include digital forensics, data hiding, image processing, and pattern recognition.

XINTAO DUAN received the Ph.D. degree from Shanghai University, Shanghai, China, in 2011. He is currently an Associate Professor with the College of Computer and Information Engineering, Henan Normal University. His major research interests include image processing, deep learning, and information security.

LIPING GAO received the B.Sc. and master's degree in computer science from Shandong Normal University, China, in 2002 and 2005, respectively, and the Ph.D. degree in computer science from Fudan University, China, in 2009. She is doing her research work in the University of Shanghai for Science and Technology as an Associate Professor. Her current research interests include CSCW, heterogeneous collaboration, and collaborative engineering.