IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received February 26, 2019, accepted April 3, 2019, date of publication April 15, 2019, date of current version April 29, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2911207

Robust Image Hashing Scheme Based on Low-Rank Decomposition and Path Integral LBP

HENGFU YANG[1,2], **JIANPING YIN**[3], **AND YING YANG**[4,5]

[1]Department of Information Science and Engineering, Hunan First Normal University, Changsha 410205, China
[2]School of Computer, National University of Defense Technology, Changsha 410073, China
[3]School of Computer Science and Network Security, Dongguan University of Technology, Dongguan 523808, China
[4]Amazon Web Service, Inc., Seattle, WA 98108-1226, USA
[5]School of Computer and Control Engineering, Minjiang University, Fuzhou 350108, China

Corresponding authors: Hengfu Yang (hengfuyang@hotmail.com) and Jianping Yin (jpyin@dgut.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 61872408 and Grant 61672528, in part by the Social Science Fund of Hunan Province under Grant 16YBA102, and in part by the Research Fund of the Hunan Provincial Key Laboratory of Informationization Technology for Basic Education under Grant 2015TP1017.

ABSTRACT This paper presents a robust image hashing algorithm that exploits low-rank decomposition and path integral local binary pattern (pi-LBP), referred to LRPL hashing. The proposed algorithm generates a compact binary sequence from a low-rank component of the normalized image as the hash code. Considering the excellent texture structure description ability of pi-LBP features, the new hashing algorithm extracts a feature vector from a low-rank feature matrix with pi-LBP. The pi-LBP feature vector is then encrypted by using the logistic map to produce the final hash sequence. The Hamming distance between the hash sequences is employed to authenticate the tested image. The experimental results demonstrate that the proposed LRPL hashing method has better robustness and discrimination compared with existing hashing algorithms.

INDEX TERMS Image hashing, low-rank decomposition, path integral local binary pattern.

I. INTRODUCTION

Robust image hashing is a technique generating a compact digest of an input image for representing image contents. The hash values are computed on the basis of perceptual contents with the ability of tolerating content-preserving distortions [1], [2]. A robust image hashing function produces similar hash values for images with same visual appearance but also is sensitive to content-changing distortions and malicious attacks. Robust image hashing received extensive attention in recent decades [3], [4]. Many image hashing algorithms are proposed and widely used in many fields including image authentication, digital watermarking, image retrieval, image copy detection, image quality assessment, and multimedia forensics [5]–[10].

As we know, feature extraction is a key procedure of image hash generation. According to feature extraction methods, robust image hashing algorithms can be roughly classified into five categories: transform based image hashing algorithms, image hashing using low-level features, image hashing based on moment invariants, image hashing based on matrix decomposition, and other image hashing methods.

A. TRANSFORM BASED IMAGE HASHING

Since transform coefficients are more convenient to extract image features than spatial coefficients, transform based image hashing methods has better robustness compared with some spatial image hashing ones. Tang *et al.* noticed that block entropies are approximately linearly changed by content-preserving manipulations and presented a hashing algorithm based on block entropies and Discrete Wavelet Transform (DWT) [11]. The hashing algorithm extracts entropies of image sub-blocks and uses a single-level 2D DWT to compress feature vectors. Correlation coefficient is employed to evaluate similarity between two different hashes. Experimental results show that the proposed image hashing scheme can resist content-preserving operations. Liu *et al.* proposed a robust image hashing method based on radon transform [4]. In this image hashing algorithm, the mapping coefficient matrix of the circular areas surrounding feature points is first obtained by Radon transform, and then the

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

invariant moments are computed from the coefficient matrix. The hashing value is generated by considering both feature vectors and invariant moments. Ouyang *et al.* employs the quaternion discrete Fourier transform (QDFT) and the log-polar transform to design a novel robust image hashing scheme for image authentication [12]. The presented hashing method generates a secondary image using a log-polar transform, and extracts features from low frequency QDFT coefficients and produces the final hash sequence according to the correlation of these QDFT coefficients. The generated image hash is sensitive to image content alterations and robust to the content-preserving operations.

B. IMAGE HASHING USING LOW-LEVEL FEATURES

Local feature points have been widely used in computer vision fields, such as image retrieval, pattern matching and object detection. Moreover, image feature points are also important to robust image hash generation. Lv and Wang use local feature points to design a shape-contexts-based image hashing method. In this approach, a robust SIFT-Harris detector is used to select the most stable SIFT key-points against content-preserving distortions, and the robust image hash sequence is generated by embedding the local features into shape-contexts-based descriptors [13]. Robust salient key-points detection enhanced robustness of the image hashing. Yan *et al.* proposed a multi-scale image hashing method using adaptive local features [14]. It gets the location-context information of the features using adaptive and local feature extraction techniques. The final hash is inserted into the image before transmission. Image authentication can be achieved by comparing the inserted hash code with the extracted hash code. Davarzani *et al.* [15] proposed a robust image hashing algorithm using center-symmetric local binary patterns (CSLBP) which extracts CSLBP features from each non-overlapping block of the input image and apply inner product of the CSLBP feature vector and a pseudorandom weight vector for each block to generate the final hash code. Experimental results show that the proposed image hashing algorithm is robust to common signal processing attacks. Qin *et al.* designed a robust image hashing scheme using perceptual texture and structure features [16]. First it employs the dual-cross pattern (DCP) to produce two coded maps representing textural information in horizontal-vertical and diagonal directions, respectively, and then uses histogram technique to extract the DCP-based textural features. At the same time, salient structural features are extracted from the frequency coefficients and selective-sampled blocks containing the richest corner points. Finally, the hash sequence is derived from the two types of extracted features.

C. IMAGE HASHING BASED ON MOMENT INVARIANTS

Zhao and Wei proposed an image hashing method based on Zernike moments [17]. The hashing algorithm calculates Zernike moments of image sub-blocks to form the intermediate hash, and conducts pseudo-random permutation on the intermediate hash to generate the final hash sequence. This method can be used to detect and locate image tampering. Tang *et al.* presented a perceptual hashing for color images. The input RGB image is firstly converted to HSI color space, and then invariant moments from both RGB and HSI color spaces are calculated. Finally, these invariant moments is concatenated to form an image hash. Experiments show good robustness and discriminative capability of Tang *et al.*'s method [18]. Chen *et al.* developed an image hashing scheme based on invariants of radial Tchebichef moments [19]. The hash values are produced by adaptive quantization of the invariants of radial Tchebichef moments. Since the radial Tchebichef moments have good orthogonality and robustness, the proposed hashing algorithm is able to achieve satisfactory robustness and discrimination. Ouyang *et al.* proposed a robust image hashing method using quaternion Zernike moments (QZMs) [20]. QZMs are suitable to deal with three channels of color images, so the hash value generated by QZM has good compactness and robustness. In a word, moments-based image hashing schemes achieve good robustness, but it is hard for them to resist shear attacks.

D. IMAGE HASHING BASED ON MATRIX DECOMPOSITION

Since non-negative matrix factorizations (NMFs) have the additivity property capturing local components of the image, and the manifestation of independent identically distributed noise on NMF vectors for spatial domain geometric attacks, Monga and Mihcak proposed a robust image hashing scheme based on NMF [21]. To enhance the robustness, in 2014, Tang *et al.* [22] combined a ring partition with NMF and devised a rotation against image hashing method. In this approach, a ring partition is used to construct the rotation-invariant secondary image, final hash sequence is obtained from NMF coefficients. Experiments show that the proposed hashing is robust against common content-preserving operations, such as image rotation, JPEG compression, low-pass filtering, brightness adjustment. Ghouti [23] proposed a perceptual color image hashing using the quaternion singular value decomposition (Q-SVD). This algorithm uses quaternion representations to handle the RGB components, and employs Q-SVD decomposition to obtain the low-rank approximation of quaternion matrices. The generated hash code can resist geometric attacks. Because Matrix decomposition provides low-rank approximation of digital images and reduces feature dimension, matrix decomposition-based image hashing algorithms can generate compact and robust hash sequence.

E. OTHER IMAGE HASHING METHODS

Monga *et al.* [24] presented a perceptual image hashing using clustering. This hashing method uses perceptually significant feature extractor to produce intermediate hash and then generates final hash by data clustering. Naoe and Takefuji exploit neural network to generate image hashing which does not embed any data into the content but is able to extract meaningful data from target image [25]. The output signal from the trained neural network is used as image hash value. In view of the extraordinary image representation learning ability of deep neural networks, Zhu *et al.* [26] designed a hashing method using deep neural network which gains improvement over several existing image hashing methods. In order to enhance the robustness against rotation manipulations, Li *et al.* proposed a robust-hash function based on random Gabor filtering and dithered lattice vector quantization (LVQ) [27]. In this hashing scheme, the conventional Gabor filter is adapted to be rotation invariant, and the rotation-invariant filter is randomized to facilitate secure feature extraction.

Recently, Low-rank analysis is used in perceptual image hashing because of the low-rank property of natural images. Liu *et al.* [28] proposed an image hashing scheme based on Low-Rank and Sparse Representation using Low-Rank Representation (LRR) to get image feature matrix and error matrix. Then the properties of dimension reduction and tampering recovery inherent in LLR and Compressive Sensing are used to recover primary feature. Finally, Compressive Sensing is used to encrypt and compress the feature vector for the hash generation. In 2015, Li and Wang proposed low-rank and sparse decomposition-based image hashing algorithm [29], which combines compressive sampling and random projection to aggregate the low-rank approximation of input image and the spatial layout of salient components into final binary hash and can achieve high robustness and discriminability. In addition, Local Binary Pattern (LBP) is a simple effective local texture feature and widely used in the fields of face analysis, image segmentation, image retrieval and image security et al [30]–[36]. Some image hashing algorithms are able to generate hash sequence from LBP features [15], [37]–[39].

In order to enhance the robustness of the image hash sequence, we design a new robust image hashing algorithm using low-rank decomposition and path integral LBP (pi-LBP). The low-rank decomposition is employed to obtain the low-rank approximation of the input image, and pi-LBP features are extracted so as to enhance the robustness of the generated hash value. The rest of the paper is organized as follows. Section 2 provides a detailed account of the proposed image hashing scheme. Experiments and analysis are given in Section 3. Finally, in Section 4 we draw conclusions and briefly discuss some possible future work with respect to robust image hashing.

II. THE PROPOSED LRPL HASHING METHOD

In this scheme, image hash generation is viewed as compact feature extraction from low-rank approximation of the input image. The whole robust image hashing generation includes three steps, preprocessing, low-rank decomposition pi-LBP feature extraction and hash generation, as shown in Figure 1.

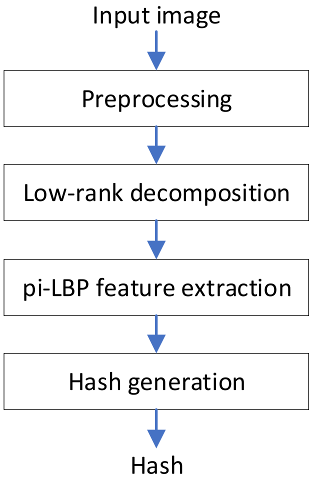


FIGURE 1. Block diagram of the proposed image hashing.

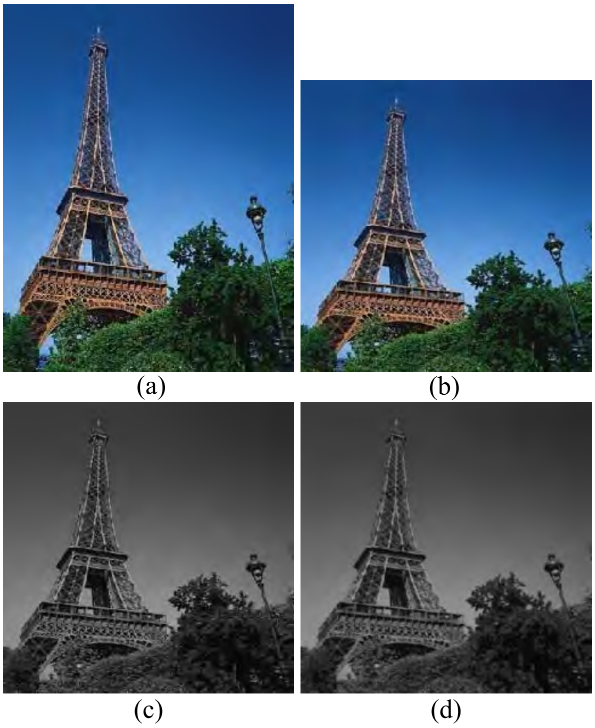


FIGURE 2. An example of preprocessing. (a) Input image. (b) Resized image. (c) L* component. (d) Filtered image.

A. PREPROCESSING

To eliminate the influence of common content preserving operations on the input image, a preprocessing is indispensable for hash generation, which includes interpolation, color space conversion and low-pass filtering. Firstly, the bi-linear interpolation is used to resize the input image to a fixed size M × M, so as to resist image rescaling operations. Then, for RGB image, the resized image is converted into CIE L*a*b* color space since L* component closely matches human perception of lightness. Finally, a Gaussian low-pass filtering is applied on the L* component to alleviate influence of minor changes of the input image. Figure 2 illustrates an example of the preprocessing.

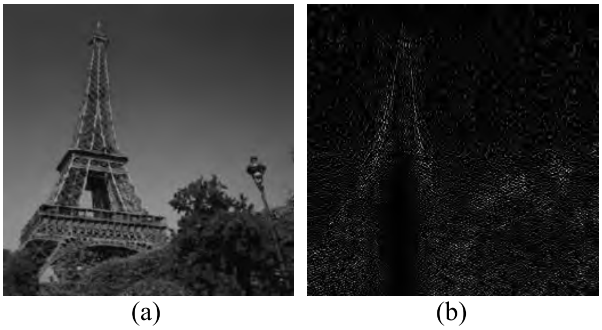


FIGURE 3. Low-rank representation of a test image. (a) Low-rank component. (b) Sparse component.

B. LOW-RANK DECOMPOSITION

Let X be the secondary image after preprocessing as shown in Fig.2 (c). Considering the consistency between the non-local self-similarity and low-rank property in digital images, we apply LRR to the secondary image X and obtain the low-rank approximation of the secondary image. So that the influence of unstable patterns in the input image can be avoided since they may not survive content preserving operations. The secondary image is decomposed into a sparse component and a low-rank component. This can be written as follows.

$$\begin{aligned} &\min_{Z,E} \ \|Z\|_* + \lambda \|E\|_{2,1} \\ &s.t.\ X = XZ + E \end{aligned} \tag{1}$$

where low-rank matrix Z represents the principle structures of the secondary image X, and E is a sparse matrix indicating salient components. $\|\bullet\|_*$ denotes the matrix nuclear norm, $\|\bullet\|_{2,1}$ is the $l_{2,1}$ norm defined as the sum of l_2 norms of the column of matrix E, and λ is a parameter which controls the importance of the sparsity error term E.

Through LRR, the principle features of the input image can be obtained to form image hash, it is beneficial to enhancing the robustness of the image hash. We solve (1) by the inexact augmented Lagrange multipliers (IALM) algorithm [40]. Fig.3 shows the decomposition results of a test image. The low-rank component and sparse component of the secondary image X are illustrated in Fig.3 (a) and Fig.3 (b), respectively.

C. PI-LBP FEATURE EXTRACTION

Feature extraction is a key step in all perceptual image hashing schemes. Robust features will lead to perceptual robustness of image hashing, so robust feature extraction is the main challenge in image hashing schemes. Local binary pattern (LBP) is a simple and effective tool for texture analysis. LBP and its variants have been widely applied in various computer vision applications such as image retrieval, image segmentation and image recognition. Because path integral LBP (pi-LBP) [41] can achieve better texture classification performance, pi-LBP is used to extract image features from the secondary image during the image hash generation. The pi-LBP calculate feature vector according to the different scale pixels along a specific path as illustrated in Fig. 4.

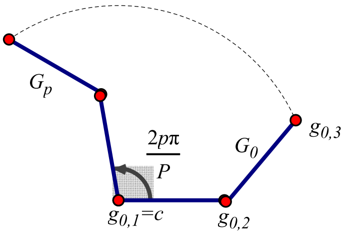


FIGURE 4. Illustration of the pi-LBP operator.

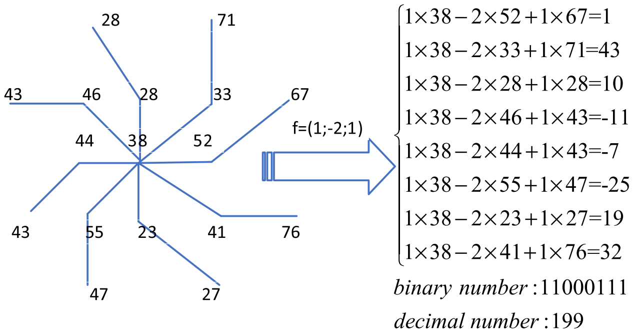


FIGURE 5. The computing of pi-LBP feature.

Given that $G_p = (g_{p,1}, \ldots, g_{p,k})$, $p = 0, \ldots, P-1$ is a specific path, where P is the number of neighbor pixels. The elements of $g_{p,1}, \ldots, g_{p,k}$ of G_p are the pixels starting with $g_{0,1} = c$. In fact, each path G_p can be obtained by rotating G_0 counterclockwise with $2p\pi/P$ degree. The pi-LBP feature can be computed by

$$pi - LBP_{P,G_0,f} = \sum_{p=0}^{P-1} s\left(\sum_{i=1}^{k} f(i)\, g_{p,i}\right) 2^p \tag{2}$$

where $f = (f(1), f(2), \cdots, f(k))$ is a filter satisfying $\sum_{i=1}^{k} f(i) = 0$.

An example of pi-LBP computing is further shown in Fig. 5. Assume that P = 8, G_0 is a path containing three pixels, and the filter f is (1;-2;1). The pi-LBP will have value 199 according to the Eq.(1).

Finally, in order to resist rotation attacks, we extract rotation invariant pi-LBP feature for perceptual image hash, the Rotation pi-LBP feature descriptor is defined as follows.

$$pi - LBP^{riu}_{P,G_0,f} = \begin{cases} s\left(\sum_{i=1}^{k} f(i)\, g_{p,i}\right) & if\ U\left(pi - LBP^{riu}_{P,G_0,f}\right) \le 2 \\ P+1 & otherwise \end{cases} \tag{3}$$

where function U returns the number of bitwise transitions from 0 to 1 or 1 to 0 in the binary form.

D. HASH GENERATION AND HASH AUTHENTICATION

Given the image sub-block size $l \times l$ for pi-LBP feature extraction, the histogram of each block is computed to form the Histogram vector H with size $N = (P+2)\left(\frac{M}{l}\right)^2$ since each histogram has P + 2 bins. To obtain compact hash representation, the principal component analysis (PCA) is used to produce compact feature from histogram feature H. it can be written as,

$$\hat{H} = \left[\hat{H}_1, \cdots, \hat{H}_q\right], \quad q \ll N \tag{4}$$

Then apply a zero-mean normalization to histogram feature $\hat{H}$ and obtain a normalized histogram feature $\bar{H}$. Subsequently, a binarization operation is performed to get the binary pi-LBP histogram feature sequence H'.

To enhance the security of the image hashing scheme, logistic map is used to generate a binary chaotic sequence S [42]. Finally, final image hash $\tilde{H}$ is generated using an exclusive operation of the pi-LBP binary sequence H' and the binary chaotic sequence S.

$$\tilde{H} = H' \otimes S \tag{5}$$

where $\otimes$ denotes bitwise exclusive operator.

The normalized Hamming distance was usually adopted to measure the similarity between two hashes, which can be defined as:

$$D_H\left(\tilde{H}_1, \tilde{H}_2\right) = \frac{1}{q}\sum_{i=1}^{q}\left|\tilde{H}_1(i) - \tilde{H}_2(i)\right| \tag{6}$$

Usually, smaller Hamming distance means more similar images. Given threshold τ, if the Hamming distance is greater than threshold τ, the images will be judged as visually identical image. So, we can authenticate whether a test image is authentic using the following equation.

$$\begin{cases} \textit{visually identical}, & \textit{if } D_H < \tau \\ \textit{tampered or distinct}, & \textit{else} \end{cases} \tag{7}$$

III. EXPERIMENTAL RESULTS

To validate the robustness and discrimination of our image hashing algorithm, the proposed image hashing algorithm is conducted on Caltech-256 image database supported by California Institute of Technology [43]. It consists of 257 categories and includes 30607 images with various different sizes. To obtain visually similar image, some commonly used content preserving operations are selected as shown in Table 1. Fig. 6 shows some typical images from the image dataset. In the experiments, the input image is resize to a standard size 256 × 256; a 3 × 3 Gaussian low-pass filter with zero mean and one unit standard deviation is taken; the block size is 5 × 5 during LBP feature extraction, neighbor pixel number P is 8, and the final hash length q = 300.

TABLE 1. Content-preserving operations for robustness testing.

Attacks	Parameters
JPEG compression	Quality factor ∈[30, 100]
3×3 Gaussian filtering	Standard deviation 0.3,0.4,…,1.0
median filtering	Filter size ∈[3, 15]
Pepper & salt noise	noise density ∈[1%,10%]
Image scaling	Scaling ratio∈[0.5, 2.0]
Gamma correction	γ=0.7,0.9,1.1,1.2
Image rotation	Rotation angle∈[0, 3]

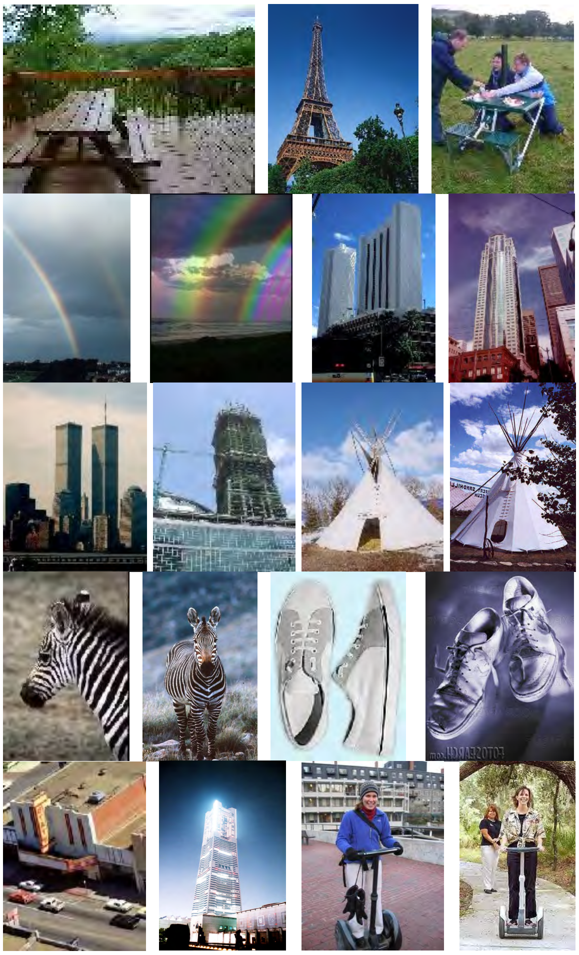


FIGURE 6. Some test images.

A. THRESHOLD SELECTION

We select 10 different color images from each category of the Caltech-256 image database as test images including 257*10 = 2570 different images. After conducting the content-preserving operations in Table 1, there are 45 visually similar images for each test image. Therefore, we can obtain 2570 × 45 = 115650 pairs of visually similar images for all 2570 test images. In addition, note that there are 2570 × (2570–1)/2 = 3301165 pairs of visually distinct images.

By applying our LRPL hashing algorithm to all test images and their visually similar versions, we can obtain 115650 and 3301165 pairs of hashes for visually similar and distinct image pairs, respectively. Thus, there are totally 115650 and 3301165 Hamming distances that can be obtained for all visually similar and distinct images pairs, respectively.

TABLE 2. Collision probability under various thresholds.

Threshold	Collision probability				
	LRPL	Ouyang et al.'s scheme [12]	Qin et al.'s scheme [16]	Chen et al.'s scheme [19]	Liu et al.’s scheme [28]
0.26	9.7993×10^{-6}	4.2906×10^{-4}	3.6293×10^{-4}	2.0518×10^{-4}	0.3694
0.24	1.7458×10^{-6}	1.1471×10^{-4}	2.5116×10^{-4}	4.7366×10^{-5}	0.1351
0.22	2.7322×10^{-7}	2.7291×10^{-5}	1.7218×10^{-4}	9.5967×10^{-6}	0.0306
0.20	3.7465×10^{-8}	5.7736×10^{-6}	1.1693×10^{-4}	1.7050×10^{-6}	0.0041
0.18	4.4989×10^{-9}	1.0854×10^{-6}	7.8660×10^{-5}	2.6542×10^{-7}	3.2451×10^{-4}
0.16	4.7290×10^{-10}	1.8120×10^{-7}	5.2417×10^{-5}	3.6185×10^{-8}	1.4608×10^{-5}
0.14	4.3496×10^{-11}	2.6851×10^{-8}	3.4600×10^{-5}	4.3179×10^{-9}	3.7352×10^{-7}
0.12	3.4996×10^{-12}	3.5304×10^{-9}	$2.2623\ \times10^{-5}$	4.5079×10^{-10}	5.3909×10^{-9}
0.10	2.3624×10^{-13}	4.1170×10^{-10}	1.4651×10^{-5}	4.1161×10^{-11}	4.3729×10^{-11}

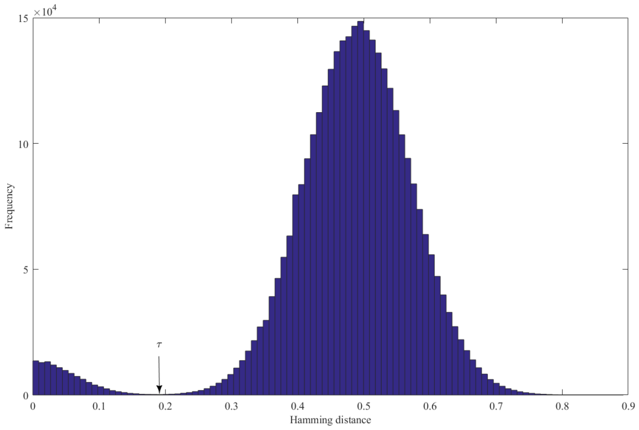


FIGURE 7. Distribution of normalized hamming distances.

Fig. 7 shows the distribution of the normalized Hamming distances for all visually similar and distinct images pairs.

Given the threshold t, the Hamming distances are classified into two classes C_1 and C_2, C_1 denotes Hamming distances with levels $[0, \ldots, t]$, and C_2 denotes Hamming distances with levels $(t, \ldots, 1]$. So, Class C_1 and class C_2 correspond to visually similar image pairs and visually distinct image pairs, respectively. The optimal threshold τ can be obtained by the maximizing inter-class variance method [44]. Let $\omega_1(t)$, $\omega_2(t)$ be the probabilities of class occurrence, and $\mu_1(t)$, $\mu_2(t)$ be the class mean levels, then the inter-class variance $Va(t)$ can be written as

$$Va(t) = \omega_1(t)\,\omega_2(t)\,(u_1(t) - u_2(t))^2 \tag{8}$$

It can be easily verified that the following equation holds for any choice of t.

$$\omega_1(t) + \omega_2(t) = 1 \tag{9}$$

Thus, the optimal threshold τ that maximizes inter-class variance $Va(t)$ can be is computed by sequential search in Eq.(7). Namely, the optimal threshold is

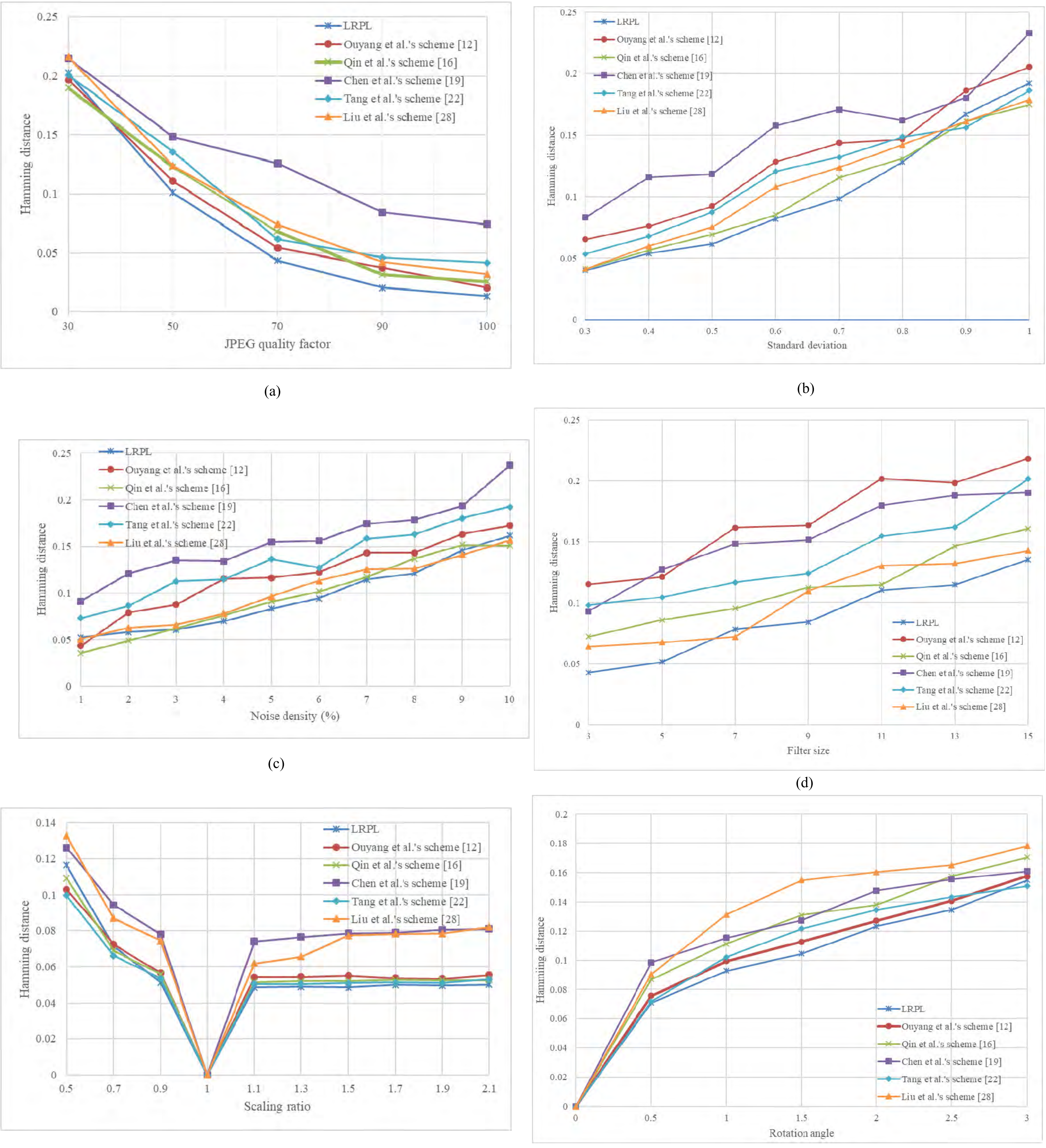
$$\tau = argmax\,(Va(t)) = 0.1926 \tag{10}$$

B. ROBUSTNESS

In order to evaluate the robustness of the robust image hashing scheme, we calculate perceptual image hash values for all 115650 pairs of visually similar images, and then their similarity is evaluated with normalized Hamming distance. In addition, we compared our hashing algorithm with previous hashing methods, i.e., Ouyang *et al.*’s scheme [12], Qin *et al.*’s scheme [16], Chen *et al.*’s scheme [19], Tang *et al.*’s scheme [22] and Liu *et al.*’s hashing scheme [28]. Fig.8 shows the perceptual robustness comparison results for our hashing scheme and the other four schemes. As to the proposed hashing scheme, it can be observed that the mean Hamming distances are less than $\tau = 0.1926$ for all the content-preserving operations. In general, the mean hamming distances of proposed LRPL hashing method are less than those of the five schemes [12], [16], [19], [22], [28]. Therefore, it can be concluded that our hashing scheme has better robustness against common content-preserving operations compared to the previous schemes [12], [16], [19], [22], [28].

C. DISCRIMINATION

A test image dataset with 2570 images built by choosing 10 different images from each category of the Caltech-256 image database is used to evaluate discrimination capacity of the proposed image hashing method, and there are 3301165 pairs of visually distinct images. Thus, there are 3301165 normalized Hamming distances produced between the hash pairs of different images. It can be observed that the distribution of normalized Hamming distance proximately obeys a normal distribution with mean $\mu = 0.491$ and standard variation $\delta = 0.0541$. So, given the threshold, the collision probability can be computed. Table 2 shows the collision probabilities under different thresholds.



(a)

(b)

(c)

(d)

(e)

(f)

FIGURE 8. Robustness comparison results among different image hashing methods based on Caltech-256 dataset. (a) JPEG compression. (b) Gaussian filtering. (c) Pepper & salt noise. (d) Median filtering. (e) Image scaling. (f) Image rotation.

It can be concluded that a small threshold will result in a low collision probability, i.e. good discrimination. When compared with Ouyang *et al.*'s scheme [12], Qin *et al.*'s scheme [16], Chen *et al.*'s scheme [19] and Liu *et al.*'s hashing scheme [28], LRPL scheme has lower collision probability. This indicates that proposed LRPL image hashing algorithm has better discrimination ability.

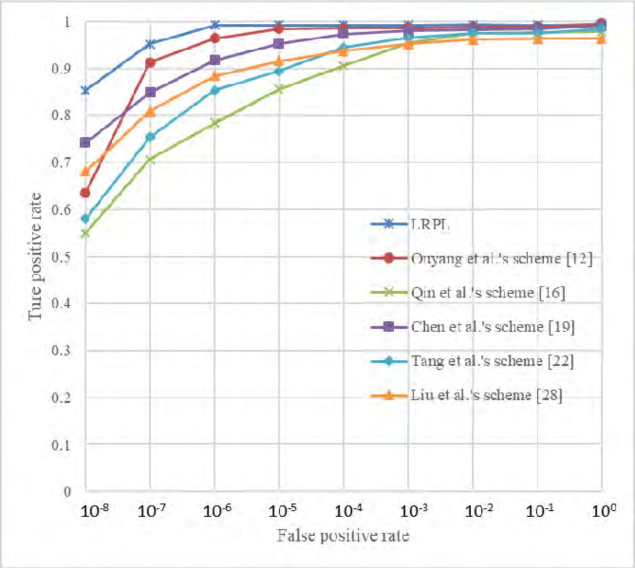


FIGURE 9. ROC curve comparison among different hashing methods.

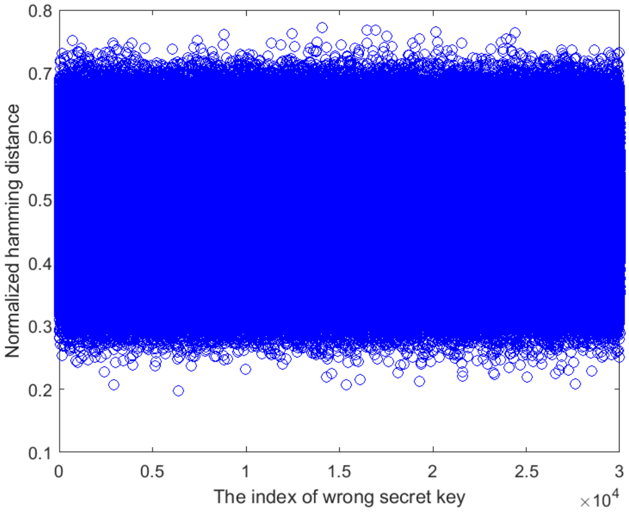


FIGURE 10. Distribution of Hamming distances between hash pairs with the correct secret key and 30000 wrong secret keys.

D. PERFORMANCE COMPARISON BY ROC

The receiver operating characteristic (ROC) curve is employed to evaluate the overall performance of different image hashing. Fig. 9 illustrates the ROC curve comparison among some image hashing methods such as Ouyang *et al.*'s scheme [12], Qin *et al.*'s scheme [16], Chen *et al.*'s scheme [19], Tang *et al.*'s scheme [22] and Liu *et al.*'s hashing scheme [28]. The closer the ROC curve is to the upper left corner, the better the classification capability of the image hashing scheme. It can be easily found that our LRPL hashing scheme has better ability to distinguish content-preserving operations from malicious ones than the other five methods [12], [16], [19], [22], [28].

E. SECURITY ANALYSIS

In our experiments, 30000 wrong secret keys are used to test the security of the proposed image hashing system. The distribution of Hamming distances for test image pairs between correct and wrong secret keys is shown in Fig. 10. It shows that all these Hamming distances are greater than τ (0.1926). Without the knowledge of the key, it is very difficult for the attacker to get the content and the same hash of original image. The probability of attackers getting the same hash is about 0.5^{300} since our hash length is 300 bits. This indicates that our proposed hashing scheme is highly key-dependent by the use of logistic chaotic encryption during hash generation.

IV. CONCLUSIONS

We presented a new robust image hashing algorithm based on low-rank decomposition and path integral local binary pattern (pi-LBP), achieving good tradeoff between robustness and discrimination. The main contribution are the low-rank feature extraction and the use of pi-LBP during hash generation. the low-rank feature matrix and pi-LBP feature extraction are used to ensure the good robustness and discrimination. The logistic chaotic encryption is exploited to enhance the hash security. Experimental results also demonstrate that the proposed LRPL Hashing algorithm is robust to common content-preserving operations, and outperforms some existing hashing algorithms. In future, we will develop robust image hashing algorithms for big data using low-rank representation.

REFERENCES

[1] 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.

[2] C. Qin, X. Chen, J. Dong, and X. Zhang, "Perceptual image hashing with selective sampling for salient structure features," *Displays*, vol. 45, pp. 26–37, Dec. 2016.

[3] N. D. Gharde, D. M. Thounaojam, B. Soni, and S. K. Biswas, "Robust perceptual image hashing using fuzzy color histogram," *Multimedia Tools Appl.*, vol. 77, no. 23, pp. 30815–30840, Dec. 2018.

[4] Y. Liu, G. Xin, and Y. Xiao, "Robust image hashing using radon transform and invariant features," *Radioengineering*, vol. 25, no. 3, pp. 556–564, Sep. 2016.

[5] Y. Cui, J. Jiang, Z. Lai, Z. Hu, and W. Wong, "Supervised discrete discriminant hashing for image retrieval," *Pattern Recognit.*, vol. 78, pp. 79–90, Jun. 2018.

[6] S. Liu *et al.*, "Perceptual uniform descriptor and ranking on manifold for image retrieval," *Inf. Sci.*, vol. 424, pp. 235–249, Jan. 2018.

[7] J. Shi, J. Wu, Y. Li, Q. Zhang, and S. Ying, "Histopathological image classification with color pattern random binary hashing-based PCANet and matrix-form classifier," *IEEE J. Biomed. Health Inform.*, vol. 21, no. 5, pp. 1327–1337, Sep. 2017.

[8] L. Wang, X. Jiang, S. Lian, D. Hu, and D. Ye, "Image authentication based on perceptual hash using Gabor filters," *Soft Comput.*, vol. 15, no. 3, pp. 493–504, Mar. 2011.

[9] S.-H. Lee, E.-J. Lee, and K.-R. Kwon, "Multi-scale curvature-based robust hashing for vector model retrieval and authentication," *Arabian J. Sci. Eng.*, vol. 44, no. 4, pp. 3235–3251, Aug. 2019. doi: 10.1007/s13369-018-3470-1.

[10] Z. Tang, S. Wang, X. Zhang, and W. Wei, "Structural feature-based image hashing and similarity metric for tampering detection," *Fundam. Inf.*, vol. 106, no. 1, pp. 75–91, Jan. 2011.

[11] Z. J. Tang, X. Q. Zhang, Y. M. Dai, and W. W. Lan, "Perceptual image hashing using local entropies and DWT," *J. Photograph. Sci.*, vol. 61, no. 2, pp. 241–251, Nov. 2013.

[12] 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.

[13] X. Lv and Z. J. Wang, "Perceptual image hashing based on shape contexts and local feature points," *IEEE Trans. Inf. Forensics Security*, vol. 7, no. 3, pp. 1081–1093, Jun. 2012.

[14] C.-P. Yan, C.-M. Pun, and X.-C. Yuan, "Multi-scale image hashing using adaptive local feature extraction for robust tampering detection," *Signal Process.*, vol. 121, pp. 1–16, Apr. 2016.

[15] R. Davarzani, S. Mozaffari, and K. Yaghmaie, "Perceptual image hashing using center-symmetric local binary patterns," *Multimedia Tools Appl.*, vol. 75, no. 8, pp. 4639–4667, Apr. 2016.

[16] 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.

[17] Y. Zhao and W. Wei, "Perceptual image hash for tampering detection using Zernike moments," in *Proc. IEEE Int. Conf. Prog. Inform. Comput.*, Shanghai, China, Dec. 2010, pp. 738–742.

[18] Z. Tang, Y. Dai, and X. Zhang, "Perceptual hashing for color images using invariant moments," *Appl. Math. Inf. Sci.*, vol. 6, no. 2S, pp. 643S–650S, Apr. 2012.

[19] Y. Chen, W. Yu, and J. Feng, "Robust image hashing using invariants of Tchebichef moments," *Optik-Int. J. Light Electron Opt.* vol. 125, no. 19, pp. 5582–5587, 2014.

[20] J. Ouyang, X. Wen, J. Liu, and J. Chen, "Robust hashing based on quaternion Zernike moments for image authentication," *ACM Trans. Multimedia Comput.*, vol. 12, no. 4, Nov. 2016, Art. no. 63.

[21] 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.

[22] 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.

[23] L. Ghouti, "Robust perceptual color image hashing using quaternion singular value decomposition," in *Proc. IEEE ICASSP*, Florence, Italy, May 2014, pp. 3794–3798.

[24] V. Monga, A. Banerjee, and B. L. Evans, "A clustering based approach to perceptual image hashing," *IEEE Trans. Inf. Forensics Security*, vol. 1, no. 1, pp. 68–79, Mar. 2006.

[25] K. Naoe and Y. Takefuji, "Damageless image hashing using neural network," in *Proc. IEEE Int. Conf. Soft Comput. Pattern Recognit.*, Paris, France, Dec. 2010, pp. 442–447.

[26] S. Zhu, B.-N. Kang, and D. Kim, "A deep neural network based hashing for efficient image retrieval," in *Proc. IEEE SMC*, Budapest, Hungary, Oct. 2016, pp. 2483–2488.

[27] 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.

[28] H. Liu, D. Xiao, Y. Xiao, and Y. Zhang, "Robust image hashing with tampering recovery capability via low-rank and sparse representation," *Multimedia Tools Appl.*, vol. 75, no. 13, pp. 7681–7696, Jul. 2015.

[29] Y. Li and P. Wang, "Robust image hashing based on low-rank and sparse decomposition," in *Proc. IEEE ICASSP*, Shanghai, China, Mar. 2016, pp. 2154–2158.

[30] B. Yang and S. Chen, "A comparative study on local binary pattern (LBP) based face recognition: LBP histogram versus LBP image," *Neurocomputing*, vol. 120, no. 23, pp. 365–379, Nov. 2013.

[31] J.-Y. Jung, S.-W. Kim, C.-H. Yoo, W.-J. Park, and S.-J. Ko, "LBP-ferns-based feature extraction for robust facial recognition," *IEEE Trans. Consum. Electron.*, vol. 62, no. 4, pp. 446–453, Nov. 2017.

[32] S. R. Dubey, S. K. Singh, and R. K. Singh, "Multichannel decoded local binary patterns for content-based image retrieval," *IEEE Trans. Image Process.*, vol. 25, no. 9, pp. 4018–4032, Sep. 2016.

[33] T. Lan, X. Feng, Z. Xia, S. Pan, and J. Peng, "Similar trademark image retrieval integrating LBP and convolutional neural network," in *Proc. ICIG*, Shanghai, China, 2017, pp. 231–242.

[34] F. Bianconi, R. Bello-Cerezo, and P. Napoletano, "Improved opponent color local binary patterns: An effective local image descriptor for color texture classification," *Proc. SPIE*, vol. 27, no. 1, pp. 011002-1–011002-10, Dec. 2017.

[35] L. Nanni, A. Lumini, and S. Brahnam, "Survey on LBP based texture descriptors for image classification," *Expert Syst. Appl.*, vol. 39, no. 3, pp. 3634–3641, Feb. 2012.

[36] Z. Xia, X. Ma, Z. Shen, X. Sun, N. N. Xiong, and B. Jeon, "Secure image LBP feature extraction in cloud-based smart campus," *IEEE Access*, vol. 6, pp. 30392–30401, Jun. 2018.

[37] 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.

[38] S. Q. Abbas, F. Ahmed, N. Živić, and Q. Ur-Rehman, "Perceptual image hashing using SVD based noise resistant local binary pattern," in *Proc. ICUMT*, Lisbon, Portugal, 2016, pp. 401–407.

[39] H. Yang, J. Yin, and M. Jiang, "Perceptual image hashing using latent low-rank representation and uniform LBP," *Appl. Sci.*, vol. 8, no. 2, p. 317, Feb. 2018.

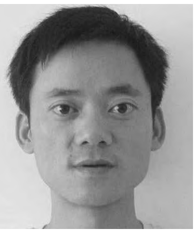
[40] C. Wang, C. Li, and J. Wang, "A modified augmented Lagrange multiplier algorithm for Toeplitz matrix completion," *Adv. Comput. Math.*, vol. 42, no. 5, pp. 1209–1224, Oct. 2016.

[41] Q. Lin and W. Qi, "Multi-scale local binary patterns based on path integral for texture classification," in *Proc. IEEE ICIP*, Quebec City, QC, Canada, Sep. 2015, pp. 26–30.

[42] Z. Tang, X. Zhang, and W. Lan, "Efficient image encryption with block shuffling and chaotic map," *Multimedia Tools Appl.*, vol. 74, no. 15, pp. 5429–5448, Aug. 2015.

[43] California Institute of Technology. (Nov. 15, 2006). *The Caltech256 Image Database*. [Online]. Available: http://www.vision.caltech.edu/Image_Datasets/Caltech256/

[44] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Trans. Syst., Man, Cybern.*, vol. 9, no. 1, pp. 62–66, Jan. 1979.



HENGFU YANG was born in 1974. He received the M.S. degree in computer application from Guizhou University, China, in 2003, and the Ph.D. degree in computer application from Hunan University, China, in 2009. Since 2015, he has been a Professor with the Department of Information Science and Engineering, Hunan First Normal University, China. He is also a Postdoctoral Associate with the School of Computer, National University of Defense Technology, China. His main research interests include image processing, information hiding, digital forensics, big data security, and multimedia encryption. He is serving as a member of the Editorial Board of the *Journal of New Media* and a Senior Member of the China Computer Federation.

JIANPING YIN was born in Hunan, China, in 1963. He received the M.Sc. and Ph.D. degrees in computer science from the National University of Defense Technology, China, in 1986 and 1990, respectively. He is currently a Professor with the School of Computer Science and Network Security, Dongguan University of Technology, China. His main research interests include information security, image processing, and pattern recognition.

YING YANG received the B.E. degree in information security and the M.E. degree in computer science and technology from the School of Computer and Communication, Hunan University, China, in 2006 and 2009, respectively, and the Ph.D. degree from the School of Engineering and Computing Sciences, Durham University, U.K., in 2013. From 2013 to 2016, he was a Postdoctoral Associate with the Computer Graphics Group, Yale University. He is currently a Manager at Amazon Web Service, Inc. He is also with the School of Computer and Control Engineering, Minjiang University, China. His research interests include digital watermarking/steganography, steganalysis, social network data analytics, and computer graphics in cultural heritage.