

Robust Hashing for Color Image Authentication Using Non-Subsampled Contourlet Transform Features and Salient Features

Qiu-Yu Zhang, Zi Yang, Qi-Yan Dou, Yan Yan

School of Computer and Communication
Lanzhou University of Technology
Gansu, Lanzhou, 730050, P. R. China

zhangqylz@163.com; 15709440859@163.com; 1147025638@qq.com; yanyan@lut.cn

Received September, 2016; revised March, 2017

ABSTRACT. A novel robust color image perceptual hashing authentication algorithm based on non-subsampled contourlet transform (NSCT) features and salient features was proposed, which can detect and locate the image tampering area more accurately. The above method simulates the human eye perception features, which extracts salient regions in a bottom-up way as local features and uses statistical properties of NSCT as global features. Global features and local features are used for image authentication. The global features are used to distinguish image whether it is similar or not, and local features are used to detect and locate tamper existing in the tested image. Experimental results show that the proposed method has good robustness and discrimination, which effectively realizes tamper detection and localization and more accurately locate.

Keywords: Image authentication, Image hashing, Non-subsampled contourlet transform (NSCT), Saliency detection, Tamper localization

1. Introduction. Image hashing approach is widely used in multimedia content authentication, database retrieval and digital watermarking. Furthermore, it has drawn great attention in recent years. When image experiences compression, filtering, geometric distortion operations, the traditional hashing based on cryptology is sensitive to the bit change of authentication information. Therefore, it is not applicable to image content authentication. So, it is significant to research an image hashing method which has robustness and distinguishes malicious tampering attacks [1].

At present, features extraction of image perceptual hashing is usually based on spatial domain and transform domain. In images, the most important part often concentrates on some key areas named salient region. By dealing with salient region, the image can be processed quickly. Features extraction is the most vital stage in existing image perceptual hashing algorithms, and different algorithms extract different features. In [2], after pseudo random blocks operation, the image was normalized to same size before conformal mapping. Then they calculated the magnitude and phase of the modified Zernike moments of the unit circles after conformal transformation. Finally, they scramble magnitude and phase of these moments to get the image hash digest. In this method, the four corners of every image block are incorporated into the hash digest, so the tamper detection performance is improved. The luminance component of color image is divided into rings, and the rings will be mapped to be a secondary image in [3]. Next, the secondary image

can be decomposed by NMF to get coefficient matrices which are used to produce the final perceptual hash digest. This method is robust to content keeping tampers, for instance, image rotation, JPEG compression, watermarking embedding and filtering. But this algorithm can only detect the tamper existing on image inscribed circle. A robust image hashing algorithm based on Zernike moments and local features was proposed [4]. The calculated five level Zernike moments are considered as global features, and the local texture features of each salient region are considered as local features. This algorithm has great discrimination which is robust to basic image processing operations. It can also detect and locate tamper in color images effectively. In [5], a color image robust perceptual hashing algorithm which uses the quaternion singular value decomposition (QSVD) was proposed. And the color correlations between the color image components are fully utilized in this algorithm. The algorithm improves robust and security for malicious attacks, but it cannot detect and locate tamper. In [6], 1-order contourlet transform was performed in every sub-block, and the image signature was generated from statistics values of transformed coefficients. By matching image signature of an original image and preset image, tamper can be located. A robust hashing method based on image structure information which can detect and locate tamper was proposed in [7]. This algorithm combines global robust hashing and local robust hashing to authenticate image. It has well robust. This method is sensitive to tamper detection after some image content keeping operations. But it is not sensitive to malicious tamper. The above researches basically have good robustness and discrimination, and can also locate and detect tamper regions. But the tamper localization is inaccurate. The tamper types cannot be clarified clearly, and the authentication data is relatively large.

In this paper, to detect and locate tamper regions in image, a novel color image perceptual hashing authentication algorithm using NSCT and salient features was proposed. The algorithm combines frequency-tuned (FT) which has fast calculation speed and good salient image effect with NSCT. The salient regions feature is extracted by FT method to detect and locate tamper, and the NSCT features are considered as global features to keep robustness and discrimination. It shortens the length of hashing digest effectively, and it can classify tampers. In addition, it can distinguish the robustness and discrimination, which can ease the contradiction between different characteristics.

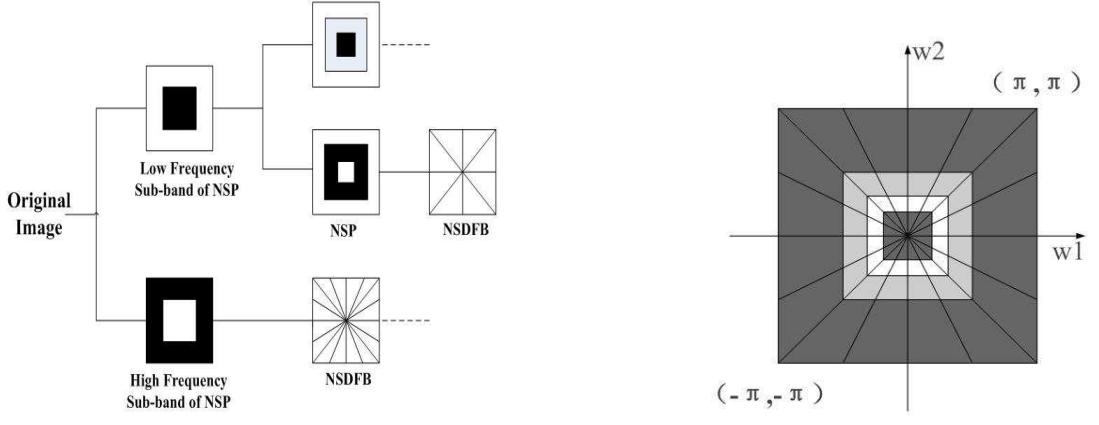
The remaining part of this paper is organized as follows. Section 2 describes theories of NSCT and image saliency detection algorithm: FT algorithm and optimizing salient regions. A detailed color image hashing algorithm is described in Section 3. Subsequently, Section 4 analyzes the experimental results and performance and compares the proposed method with other related approaches. Finally, we conclude our conclusion in Section 5.

2. Problem Statement and Preliminaries.

2.1. Non-Subsampled Contourlet Transform (NSCT). Non-Subsampled Pyramid (NSP) and Non-Subsampled Direction Filter Banks (NSDFB) make up NSCT method [8]. NSP decomposes image into multi-resolution sub images. And NSDFB can decompose image from many directions.

The detailed NSCT decomposition processes are as follows: Firstly, the image is decomposed to obtain a low frequency sub-band and a high frequency sub-band. Then bandpass sub-band image of the NSP output is transformed through NSDFB. The singular points distributed in the same direction form into a coefficient, and bandpass directional sub-band of image is obtained. If the bandpass directional sub-band coefficient image is decomposed l levels, 2^l directional sub-band images which have the same size with original image can be obtained. Thus, NSCT is anisotropic to realize the sparser image

representation. The low frequency sub-band images got from NSP are the input to next NSP. The repetition of NSP iterative operations decomposes image in multi-scale and multi-direction. Fig. 1 is the NSCT decomposition diagram.



(a) The composition structure diagram of NSCT (b) The frequency decomposition diagram of NSCT

FIGURE 1. NSCT decomposition diagram.

2.2. Detection Algorithms of Image Saliency: FT Algorithm. According to human visual selective attention mechanism, image saliency detection can search salient regions [9]. FT algorithm, proposed by Achanta [10], is a kind of salient region detection algorithm based on global comparison from the viewpoint of frequency domain. The basic idea of FT algorithm is filtering continuous frequency band from low frequency to high frequency with multiple Gaussian bandpass filters. By this way, the image is transformed into the Lab color space. And, the differences between each pixel value of filtered image and the whole image pixel average value are calculated to express saliency. The algorithm only needs to calculate the Gaussian smoothing and average, which get good results. In addition, FT algorithm calculates color distance in CIELAB color space, so it satisfies human perception. By using the proposed method, better clear segmentation boundaries can be obtained; salient regions can be evenly highlighted. Furthermore, FT has fast operating speed, and provides full resolution saliency maps.

The salient value of each pixel (x, y) in FT algorithm can be defined as follow.

$$S(x, y) = \|I_\mu - I_{whc}(x, y)\| \quad (1)$$

where I is in Lab color space, I_μ is the mean image feature vector, $I_{whc}(x, y)$ is the corresponding image pixel vector value in the Gaussian blurred version of the original image, and $\| * \|$ is Euclidean distance. In the Lab color space, each pixel location is an $[L, a, b]^T$ vector. L represents luminance, a represents the range from carmine to green and b represents the range from yellow to blue. The Eq. (1) illustrates that the Euclidean distance between mean vector of input image and Gaussian filter vector is the salient image. This way can get a salient image which removes some high frequency noises that can affect the calculation of salient value.

Several Gaussian filters make up difference of Gaussian (DoG). DoG can calculate original image after Gaussian filter. The combined DoG filters can calculate the information between the lowest frequency and the highest frequency, which calculates saliency. Each DoG is a simple band-pass filter. The DoG filter is given by

$$DoG(x, y) = \frac{1}{2\pi} \left[\frac{1}{\sigma_1^2} e^{-\frac{x^2+y^2}{2\sigma_1^2}} - \frac{1}{\sigma_2^2} e^{-\frac{x^2+y^2}{2\sigma_2^2}} \right] = G(x, y, \sigma_1) - G(x, y, \sigma_2) \quad (2)$$

2.3. Optimizing Salient Regions. The image binarization, morphology processing and optimized processing results of the salient region are shown in Fig. 2.

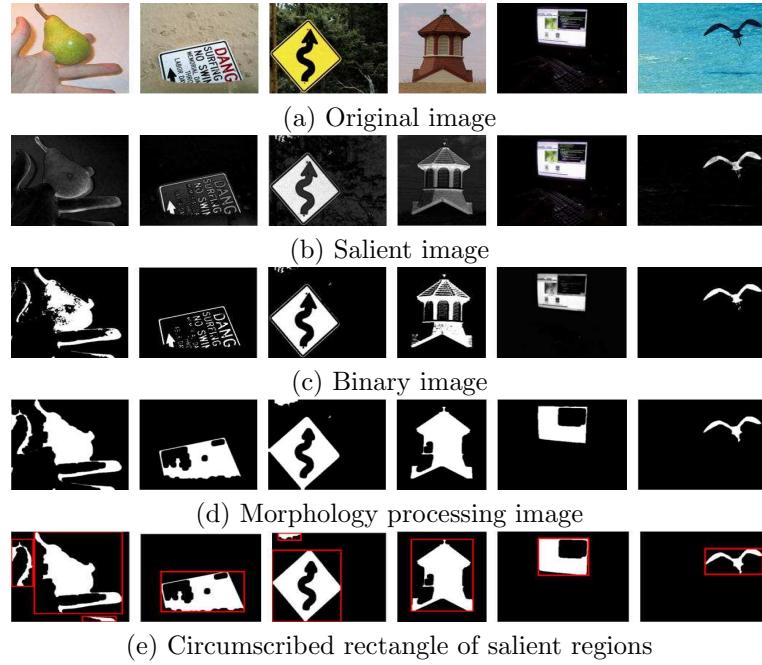


FIGURE 2. Image binarization, morphology processing, salient region.

Using FT algorithm, we can get saliency map shown in Fig. 2(b). The Otsu's method achieves binarization of saliency map, and the result is shown in Fig. 2(c). And the white part is the saliency map. Fig. 2(c) shows that the number of white areas is more and the white areas distribute disorderly. If these white regions are used to directly extract features, noise interference will be more and the length of feature is long. So the morphological processing can be adopted to optimize saliency map. In Fig. 2(c), there are some tiny area stray points, which can vary under common image processing operations. And the matching of image features is not affected, but it will lead to long hash. Therefore, the appropriate area threshold values are set. The closing operation can remove the dark and rounded corners of binary image $imgn$ and the image $imgn1$ is obtained. Then opening operation can smooth the contour of the salient object, connect the narrow gap to form an elongated curved mouth, and fill the structural elements to get $imgn2$. The image whose area is less than the given value in $imgn2$ is removed. The final salient regions are shown in Fig. 2(d). The circumscribed rectangle of each salient region is calculated to extract local features. The salient region and circumscribed rectangle are shown in Fig. 2(e).

3. The Proposed Method.

3.1. Image Hashing Extraction Method. NSTC can extract the global features. And the local features can be obtained from salient regions. These two features set make up hash sequence. The global features can meet robustness and discrimination. The local features can detect and locate tamper. Comparing with tampering location methods based on block method, salient regions approaches can shorten the length of local features effectively. And after image rotation operation, it hardly disorders image blocks which can lead to the failure of tampering location.

3.1.1. *Extraction method of global hashing.* The extraction steps are described as follows:

Step 1.: The $M \times N$ color image is resized to a standard size 512×512 , then the color space of the image is transformed from RGB to YCbCr, and the luminance component Y is taken for representation.

Step 2.: 512×512 image is divided into non-overlapping 32×32 blocks. So we have 256 blocks.

Step 3.: Take 1-order non-subsampled contourlet transform for each block, let M , N respectively represent the number of row and column of 1-order low-pass sub-band. $C_k(i, j)$ ($i, j = 1, 2, \dots, 256$) represents the coefficients of low-pass sub-band of the k -th block, F_k represents the coefficients's variance of 1th-order low-pass sub-band of the k -th block. Computing mean μ_k and variance F_k of $C_k(i, j)$

$$\mu_k = \frac{1}{M \times N} \sum_{i=1}^M \sum_{j=1}^N C_k(i, j) \quad (3)$$

$$F_k = \frac{1}{M \times N} \sum_{i=1}^M \sum_{j=1}^N (C_k(i, j) - \mu_k)^2 \quad (4)$$

The variance of every block form the intermediate hash values $H_S = (F_1, F_2, \dots, F_{256})$.

Step 4.: Quantification. The intermediate hash values $H_S = (F_1, F_2, \dots, F_{256})$ quantified using Eq. (5) to simplify hashes, where $t \in (1, 2, \dots, 255)$.

$$H_{St} = \begin{cases} 1 & F_t > F_{t+1} \\ 0 & F_t < F_{t+1} \end{cases} \quad (5)$$

The final global signature is the qualified perceptual hashing H_S .

3.1.2. *Extraction method of local hashing.* The extraction steps are described as follows:

Step 1.: 3×3 Gaussian filter can smoothing input image I so that the highest frequency can be extracted. And image processed by Gaussian filter is obtained.

Step 2.: The input image from RGB color space is converted to Lab color space. In Lab color space, three parameters L, a, b and their average value are calculated.

Step 3.: FT algorithm mentioned in Section 2.2 can get the final saliency map sm .

Step 4.: Using Otsu's method to achieve binarization of saliency map sm , then we can get binary image $imgn$. The binary images are processed by various morphological processing to obtain salient region, and the outer rectangle of the salient region is calculated. The bright mean of each circum-rectangle in original image and the location character c (the coordinate of upper left corner, length, width) of every circum-rectangle are calculated as local feature $H_Z[\mu_1, \mu_2, \dots, \mu_n, c_1, c_2, \dots, c_n]$. Where n is the number of enclosing rectangle.

The obtained image hash value is the connection between the global perceptual hash value and the local hashing value: $H_H = [H_S, H_Z]$.

3.2. Similarity Metric for Image.

3.2.1. *Similarity metric of global hashing.* Normalized Hamming distance is exploited to measure similarity of global hashing between two image hashes. Let H_{S0}, H_{S1} be two image intermediate hashes from test image and receive final hashes. Then, normalized hamming distance (The bit error rate) $D_H(\cdot)$ can be calculated by the Eq. (6).

$$D_H(H_{S0}, H_{S1}) = \frac{1}{L} \sum_{w=1}^L |H_{S0}(w) - H_{S1}(w)| \quad (6)$$

where the hash sequence length of H_{S0} and H_{S1} is L , $L=256$, w represents every bit in hashes.

3.2.2. Similarity metric of local hashing. Euclidean distance is exploited to measure similarity of local hashing. The location vector c is only as the location identification of the salient region and it doesn't participate in the calculation of similarity. If the number of element in two image's local hashing is unequal, test image is tampered. H_{Z1} is local perceptual hashing value which needs authentication and H_{Z0} is received local hashing value. If the number of two image's local hashing element is same, Euclidean distance between these two hash values is calculated. If the Euclidean distance of a block is greater than threshold T_2 , the image block is tampered. Otherwise, salient region is normal.

3.3. Image Hashing Authentication Process. H_{S1} is global hashing and H_{Z1} is local hashing. According to Eq. (6), the similarity between original and test image can be calculated. When the distance of global hashing $S_{S01} > T_1 = 0.25$, the test image is different with original image. Otherwise, when $S_{S01} < T_1$, we use local perceptual hashing to judge that the test image is similar or tampered. Comparing the number of element and the elements' value of local perceptual hashing, we can detect test image is tampered or not. In addition, the type of tamper can be judged, such as add, delete or replace content. When the content is replaced, the Euclidean distance of local perceptual hashing S_{Z01} is compared. When $S_{Z01} > T_2 = 30$, this salient region is tampered. We need to locate the tampered position.

The flow chart of image hashing authentication method based on global and local hashing is shown in Fig. 3.

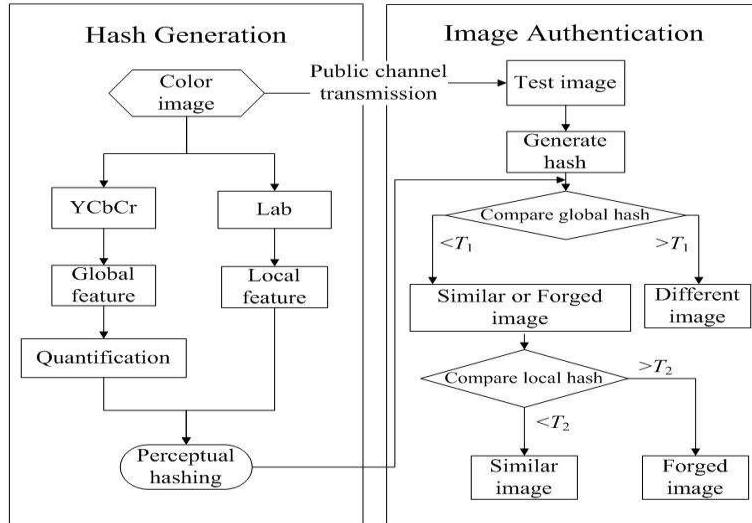


FIGURE 3. The flow of image hashing authentication.

4. Experimental Results and Performance Analysis. We choose 800 color test images from James Z. Wang Research Group Image Database [11] and 200 different images from Ground Truth Image Database [12]. The above test image data can validate robustness and uniqueness of the proposed method. In total, there are 1,000 different 24 bit true color images, includes people, food, flower, car and so on. The experiments run in Windows 7 system and the experimental platform is MATLAB 2009a. The image size range from 256×384 to 756×504 and the block size is 32×32 . Thus, we obtain 256 image blocks and the length of image hash is $L=256$. NSCT feature is global hashing in

our experiments, which can check the similarity of image. Content keeping manipulations includes gamma correction, Gaussian noise, salt and pepper noise, scaling and JPEG compression. Perceptual hashing algorithm based on global features and local features validates robustness of different intensity content keeping manipulations. In experiments, 1,000 images verify the uniqueness of our method.

We have two research objects: to emphasize the merits of FT saliency detection which is used in local hashing and to ensure the detection and localization of tamper. For these two targets, we adopt images set B database in MSRA Salient Object Database [9, 13] to validate detection and localization performances. The image library owns 5,000 images, and it includes 10 sub image sets. A part of the images set B is randomly selected to do experiment.

4.1. Robustness Test. The experiments select five images of 24 bit true color standard image: Airplane, Baboon, House, Lena and Peppers to carry out different kinds of common content keeping operations. Fig. 4 shows five standard images and Table 1 provides a specific parameters set for various content keeping operations.

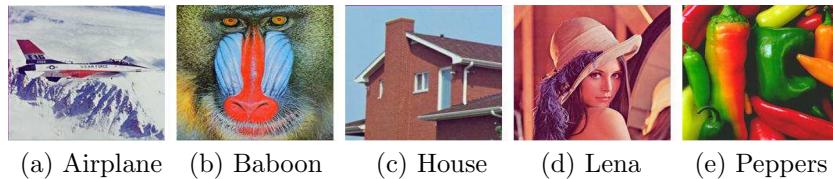


FIGURE 4. Standard color images for robustness validation.

TABLE 1. The used content-preserving operations and their parameter values

Manipulations	Parameters	Parameter values
Gamma correction	γ	0.75, 0.9, 1.1, 1.25
Scaling	Ratio	0.25, 0.5, 0.75, 1.25, 1.5, 2
JPEG compression	Quality factor	10, 20, 30, ..., 100
Brightness adjustment	Photoshop's scale	-20, -10, 10, 20
Contrast adjustment	Photoshop's scale	-20, -10, 10, 20
Gaussian noise	$\mu = 0, \sigma^2$	0.01, 0.02, 0.03, 0.05, 0.1, 0.2, 0.25, 0.3
Salt and pepper noise	Noise density	0.01, 0.02, 0.03, 0.05, 0.07, 0.1, 0.2, 0.25, 0.3
3×3 Gaussian low-pass filtering	Standard deviation	0.1, 0.2, 0.3, ..., 1.0
Rotation	Angle in degree	$\pm 315, \pm 270, \pm 225, \pm 180,$ $\pm 135, \pm 90, \pm 45, \pm 30, \pm 15, \pm 5$

Due to space constraints, standard image Baboon and Lena experiences some content keeping operations and results are shown in Fig. 5.

The original image hash is H_{S1} . The image experienced many content keeping operations produce hash H_{S2} . By Eq. (6), the Normalized Hamming distance between H_{S1} and H_{S2} can be calculated. Fig. 6 shows the Normalized Hamming distance of 5 standard test images under different common content keeping operations. In the process of rotation, images are rotated by the “loose” method. In addition to the specific angle of $-270^\circ, -180^\circ, -90^\circ, 90^\circ, 180^\circ$ and 270° , the size of image remains unchanged. Otherwise, image size increases at the other angle.

As can be seen in Fig. 6(a)-Fig. 6(i), the proposed algorithm is not only robust to gamma correction, scaling, JPEG compression, brightness adjustment, contrast adjustment, salt and pepper noise and 3×3 Gaussian low-pass filter, but also is robust to Gaussian noise. When the threshold T_1 is 0.25, the proposed algorithm is robust to all

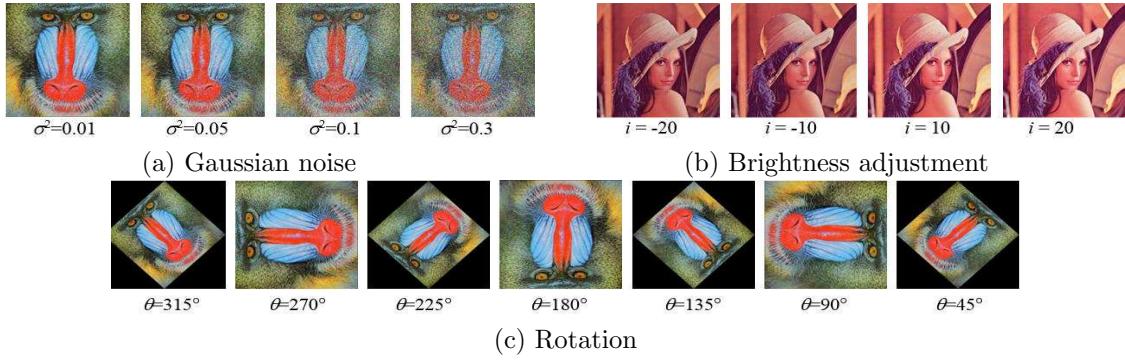


FIGURE 5. Lena and Baboon image after content-preserving operations.

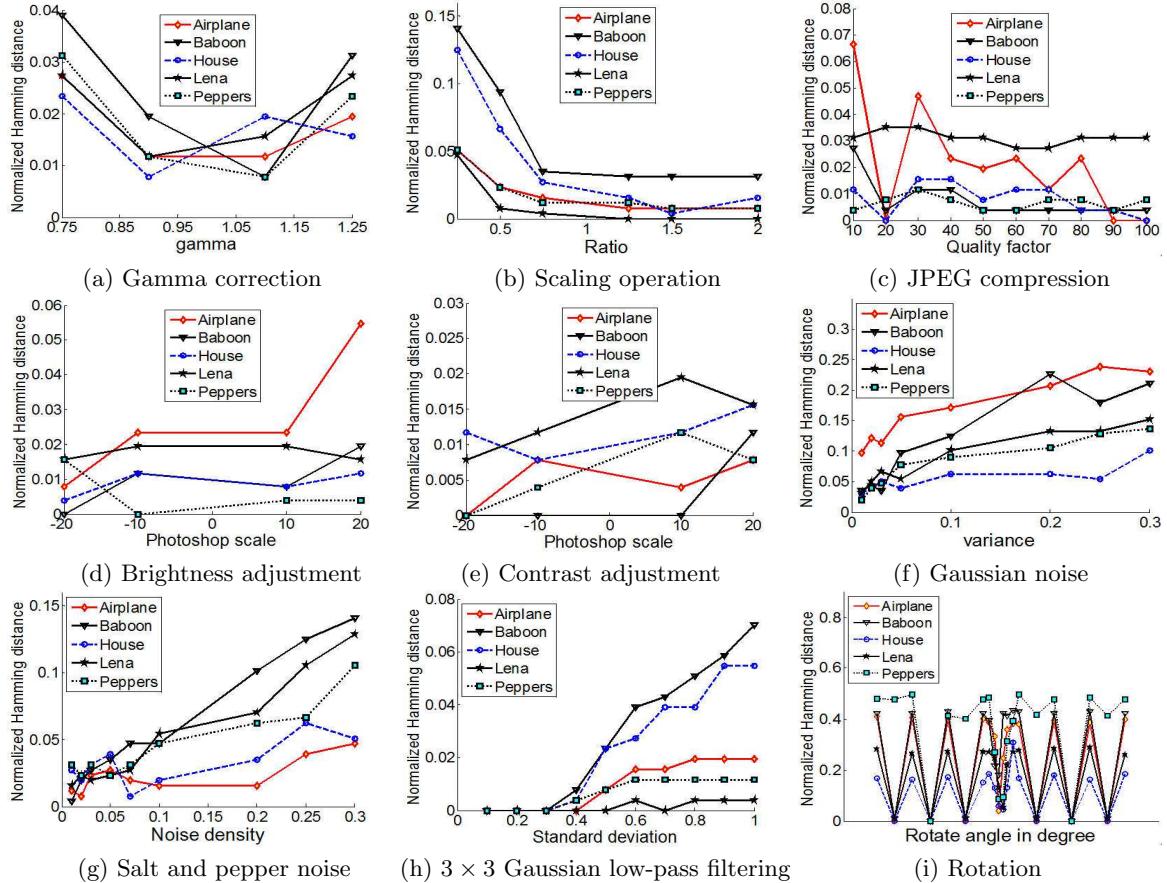


FIGURE 6. Robustness test based on five standard test images.

operations except for rotation. For the rotation operation, the Lena and House images still have good robustness, but the other 3 standard images, especially the Peppers images, are completely identified as different images. In addition, different from the Peppers images, four standard images have good robustness in the angle of $-270^\circ, -180^\circ, -90^\circ, 90^\circ, 180^\circ$ and 270° . The Normalized Hamming distance under these 6 angles are close to 0.

As can be seen in Fig. 6, except for rotation operation in Fig. 6(i), each standard test image obtains 55 similar images after 8 kinds of common content keeping operations, so we get a total of 275 images. The maximum of Normalized Hamming distance between test image and the image experienced common content keeping operations is 0.2383 and the minimum is 0. Mean distance m and standard deviation s are 0.0336 and 0.0439.

4.2. Uniqueness Analysis. The method uniqueness is same as anti-collision, and uniqueness experiment reflects the performance of anti-collision for image hashing algorithm. After applying our image hashing generation algorithm to 1,000 different color images, $C_{1000}^2=499,500$ Normalized hamming distance can be obtained. Statistical histogram represents experimental result as Fig. 7.

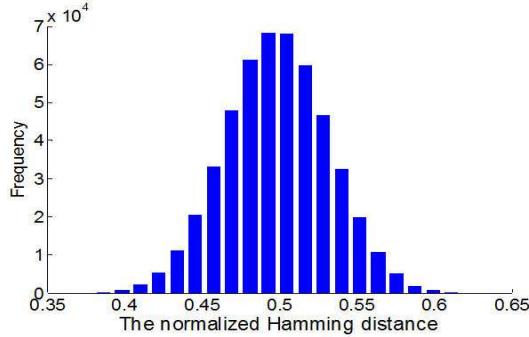


FIGURE 7. The normalized Hamming distance distribution diagram.

If normalized Hamming distance is close to 0.5, the images are different. If normalized Hamming distance is close to 0, it represents that images are similar. The distribution of normalized Hamming distance in Fig. 7 shows that distance of different image pairs mainly gathers near 0.5, and the results generally obey normal distribution. The mean μ of normalized Hamming distance among 1,000 different images is 0.4976 and the standard deviation σ is 0.0336 when threshold $T_1=0.25$.

$$P = \int_{-\infty}^T \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = 8.5926 \times 10^{-14} \quad (7)$$

According to the experimental results in Fig. 7, the collision probability is 1×10^{-14} when $T_1=0.25$, which means pretty small collision probability and good anti-collision performance.

4.3. Comparing with Robustness and Uniqueness. There are 5 standard tested images participate in robustness experiments. The images experienced content keeping operations except for rotation have 55 similar images. According to $5C_{56}^2 = 5 \times \frac{56 \times 55}{2} = 7,700$, there are altogether 7,700 similar images pairs tested results. Between pairs of similar images, the mean value of normalized Hamming distance is 0.0822. The standard deviation is 0.0825 and the minimum and maximum are 0 and 0.3867. We choose error detection rate P_E and collision rate P_C evaluate the performance of our method. The definition of P_E and P_C are Eq. (8) and Eq. (9).

$$P_E = \frac{N_E}{N_S} \quad (8)$$

$$P_C = \frac{N_C}{N_D} \quad (9)$$

where N_E is the error number of error detection of similar images and N_S is the total number of similar images. N_C is the error number of error detection of different images and N_D is the total number of different images.

Choosing different threshold, the error detection rate and collision rate can be calculated by Eq. (8) and Eq. (9). The results are shown in Table 2.

According to Table 2, the threshold is smaller, the collision rate is smaller and the error detection rate is higher. The range of threshold value T is [0.2~0.39], which keep the

TABLE 2. The relations between error detection rate, collision rate and threshold

Thresholds	P_C	P_E
0.10	1.3125×10^{-32}	0.4146
0.15	2.1991×10^{-25}	0.2054
0.20	4.1046×10^{-19}	7.67×10^{-2}
0.25	8.5926×10^{-14}	2.1×10^{-2}
0.30	2.0396×10^{-9}	4.1×10^{-3}
0.35	5.5935×10^{-6}	6×10^{-4}
0.39	6.8148×10^{-4}	1×10^{-4}
0.40	0.0018	1×10^{-4}

good robustness and uniqueness. The error detection rate and collision rate are compared with Ref. [1, 2, 16]. The relation coefficient was adopted in Ref. [1], Euclidean distance was used in Ref. [16] and the distance equation was adopted in Ref. [2] are defined as follow.

$$D(H, H') = \sqrt{\sum_{|H(k) - H'(k)| \in \Gamma_{max}} |H(k) - H'(k)|^2} \quad (10)$$

When the collision rate reaches to same order of magnitude, different distance equations match to different thresholds. As the evaluation index of image perceptual hashing method, error detection rate and collision rate are vital. Therefore, error detection rate of four approaches are shown in Table 3.

TABLE 3. Compare P_E and P_C in different methods

Hash method	Thresholds	P_E	P_C
Our algorithm	0.39	1×10^{-4}	6.8148×10^{-4}
Ref. [1]	0.6	9×10^{-4}	5×10^{-4}
Ref. [2]	40	9.93×10^{-4}	1.1×10^{-4}
Ref. [16]	3200	0.0799	1.08×10^{-4}

When the order of magnitude is 1×10^{-4} , the error detection rate of the proposed method is lower than Ref. [1, 2, 16].

4.4. Forgery Detection and Location Capability. Five standard original images, images experienced content keeping operations and images rotate 5° makes up images set. The experiment adopts Adobe Photoshop CS6 tamper 20% content of each image in images set. Then, some tampered images are shown in Fig. 8.

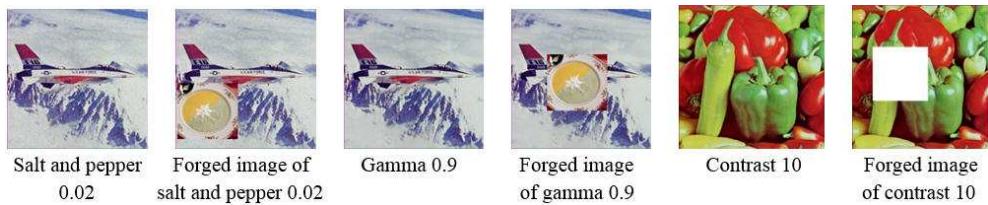


FIGURE 8. Images after common content-preserving operations and forged images.

To evaluate performance of the proposed hashing scheme, the global hashing is calculated by the method proposed in Section 3.1. Then, we calculate normal Hamming distance between tampered image and similar image. The normal Hamming distance of global hashing between original image and tampered image can be known. The maximum of normalized Hamming distance between test image and image experienced content

keeping operation is 0.2383. The min value is 0. The mean and standard deviation of test image and similar image are 0.0336 and 0.0439. Between pairs of images in 1,000 different images, the maximum of normal Hamming distance is 0.6523 and the minimum is 0.1997. The mean value is 0.4976 and standard deviation is 0.0336. Thus, the max value of normal Hamming distance between tampered and similar image's global hashing equals to 0.1680 which is smaller than the threshold 0.25 got from robustness test. Global hashing can detect images so that similar images, tampered images and different images can be judged. And we use local perceptual hashing to detect and locate tamper regions.

H_S is the global hashing value of original image. S_{S01} is the similarity distance between global hashing value of tested image and H_S . When $S_{S01} < T_1$, the local hashing value of both kinds of images can be compared. If the number of elements in local hashing H_Z of tested image is larger than original image indicates that there exists additional content in test image, the proposed method uses the position vector c (upper left coordinate, length, width) to locate the additional content. When the number of elements in local hash H_Z of tested image is smaller than original image, it indicates that the part salient content of original image was deleted, so this method can also locate the delete content. H_{Z1} is the local perceptual hash of tested images. H_{Z0} is the received local hash. When the number of elements in local hash is equal, Euclidean distance S_{Z01} is computed. When $S_{Z01} > T_2$, it indicates that the image is tampered and tamper localization is implemented.

An analysis of performance among the proposed method and Ref. [14-18] are shown in Table 4. It can be concluded that the proposed method has good overall performance, which generates hash digest with the second shortest length. After optimization of saliency, the images have less than 5 salient regions. The number of salient regions is reduced.

TABLE 4. Comparison of algorithms

	Ref. [14]	Ref. [15]	Ref. [16]	Ref. [17]	Ref. [18]	Our algorithm
Features used	Local	Local	Local	Global	Local	Global and saliency related local features
Hash length	7168 bits	250 bits	64 floating point numbers	320 bits	512 bits	256+5 × the number of salient regions bits
Robust against small-angle rotation	No	Yes	No	No	Yes	Yes
Robust against JPEG and additive noise	Yes	Yes	Yes	Yes	Yes	Yes
Robust against slight cropping	No	Yes	Yes	No	Yes	Yes
Ability to detect small area forgery	Yes	No	Yes	Yes	Yes	Yes
Ability to locate forged regions	Yes	No	No	No	Yes	Yes

To further evaluate the discrimination between similar and tampered images, the false alarm probabilities P_{ST} and missed alarm probabilities P_{TF} are defined by:

$$P_{ST} = \frac{\text{Number of natural images judged as forged images}}{\text{Total number of natural images}} \quad (11)$$

$$P_{TF} = \frac{\text{Number of forged images judged as natural images}}{\text{Total number of forged images}} \quad (12)$$

The P_{ST} and P_{TF} are applied to evaluate the detection performance of the proposed method, Monga method [16], Tang method [17] and Liu method [18]. In [14], the length of hash digest is much longer than others and its distance is a matrix. In [15], tiny

area tamper cannot be detected. The ROC curve of [14] and [15] cannot be obtained, they algorithm cannot be compared. Because of above two reasons, methods in [14] and [15] don't participate in the comparison analysis. Fig. 9 compares the ROC curves of tampering detection under rotate 5°. As the Fig. 9 shown, the proposed method has better performance than other compared algorithms under small angel rotation.

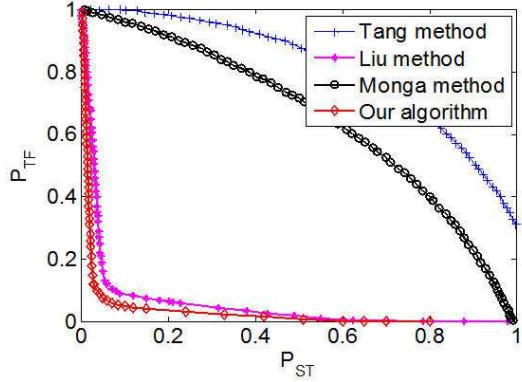


FIGURE 9. Comparison of tamper detection performance.

Table 5 shows the comparison results of the proposed method and approaches in [4] which has the same detection performance of tamper.

TABLE 5. Performance comparison

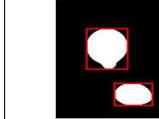
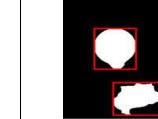
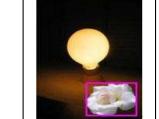
	Ref. [4]	Our algorithm
Features used	Global and saliency related local features	Global and saliency related local features
Optima thresholds	$T_1 = 7, T_2 = 50$	$T_1 = 0.25, T_2 = 30$
Hash length	560 bits	256+5 × the number of salient regions bits
Robust against small-angle rotation	Yes	Yes
Robust against JPEG coding and additive noise	Yes	Yes
Robust against slight cropping	Yes	Yes
Ability to detect small area forgery	Yes	Yes
Ability to locate forged regions	Yes	Yes

As shown in Table 5, the proposed algorithm has the same detection performance of tamper with compared method. What's more, the length of hash digest is shorter in our method. Our method has better detection performance than the approach in [4].

Some examples after tampering detection and location are shown in Table 6. The tampered images are randomly selected from Images set B of MSRA Salient Object Database using Adobe Photoshop CS6. The first column from the left is original image; the second column is tampered image. In the third column, salient regions in original image are listed. And the salient regions in tampered image are shown in the fourth column. The rectangle in the fifth column indicates the detected tampered regions. The sixth and seventh columns are image size and Euclidean distance respectively. The Euclidean distance of the same salient regions between original and tampered image in Table 6 is 0.

As shown in Table 6, the proposed method can realize the location of tamper regions in color images in different sizes and validate the type basic tampers: adding, deleting and replacing. The purple rectangles represent as the salient regions of replacing object in image; the yellow rectangles represent as the salient regions only exist in tampered image and the blue rectangles represent the salient regions exist in the original image but

TABLE 6. Local perceptual hashing distance between original and forged images

Original image	Forged image	Salient regions in original image	Salient regions in forged image	Detection result	Image size	Euclidean distance
					300×400	35.0123
					400×267	48.6593
					298×384	91.8624
					400×264	144.7009
					256×384	126.6842
					438×256	122.4091 12.6029

don't exist in the tampered image. That is to say, if there is any adding operation in tested image but doesn't exist in original one, it will be marked out by yellow rectangular in tested image; if there is any replaced content in original image, it will be marked out by purple rectangle in tested image. Moreover, the content deleted from original image will be marked out by blue rectangle in tested image. In Table 6, the first two lines are examples of replacing objects, the third and fourth lines are examples of adding objects and the last two lines are examples of deleting objects. In the first line, the egg in the lower right corner is replaced by a flower in original image; in the second line, the color of the balloon on the right is switched to another; in the third line, there add a colorful butterfly above the flower; in the fourth lines, there add a small yellow flower on the right of the original image; in the fifth image, the building in left lower corner is erased; in the sixth image, the yellow flower on the right of original image is deleted. For a reasonable range of threshold, this method chooses $T_2 = 30$ as tamper location threshold. According to experimental results of tamper detection and location, in the sixth line, there are three salient regions in original image while there are two in tampered image. The salient region on the right of tampered image turns to be much bigger after deleting the yellow flower by PS software, but the second Euclidean distance in seventh row is less than threshold T_2 which is corresponding to that region, which represents this region has not been tampered.

5. Conclusions. In this paper, we present a novel robust color image perceptual hashing authentication algorithm based on NSCT and saliency features. The image saliency detection method based on FT is adopted to extract brightness mean value of salient regions in image as local features. Then global features are the variance of NSCT coefficient for each block which comes from uniform blocks of image. The proposed method uses global features to ensure discrimination and robustness and uses local features for tamper detection and location. Experimental results show that the proposed method utilizing

global and local features to authenticate image is robust against most content keeping attacks, and it can shorten the length of hash digest effectively. In addition, it achieves well tamper detection and location, the type of tamper can also be identified.

Acknowledgment. This work is supported by the National Natural Science Foundation of China (No. 61363078), the Natural Science Foundation of Gansu Province of China (No. 1310RJYA004). The authors would like to thank the anonymous reviewers for their helpful comments and suggestions.

REFERENCES

- [1] Z. Tang, L. Ruan, C. Qin, et al, Robust image hashing with embedding vector variance of LLE, *Digital Signal Processing*, vol. 43, pp. 17–27, 2015.
- [2] Y. Zhao, S. Z. Wang, H. Yao, et al, Perceptual image hashing based on conformal mapping and Zernike moments, *Journal of Applied Sciences*, vol. 30, no. 1, pp. 75–81, 2012.
- [3] Z. Tang, X. Zhang, and S. Zhang, Robust perceptual image hashing based on ring partition and NMF, *IEEE Transactions on Knowledge and Data Engineering*, vol. 26, no. 3, pp. 711–724, 2014.
- [4] Y. Zhao, S. Wang, X. Zhang, et al, Robust hashing for image authentication using Zernike moments and local features, *IEEE Transactions on Information Forensics and Security*, vol. 8, no. 1, pp. 55–63, 2013.
- [5] L. Ghouti, Robust perceptual color image hashing using quaternion singular value decomposition, *Proc. of the 2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, pp. 3794–3798, 2014.
- [6] X. Wang, L. Zhang, and K. Pang, A Contourlet-Domain Image Signature for Content Authentication, *Proc. of the 2013 Ninth International Conference on Intelligent Information Hiding and Multimedia Signal Processing*, IEEE, pp. 555–558, 2013.
- [7] Z. Liu, Q. Li, H. Zhang, et al, An image structure information based robust hash for tamper detection and localization, *Proc. of the 2010 Sixth International Conference on Intelligent Information Hiding and Multimedia Signal Processing (IHH-MSP)*, IEEE, pp. 430–433, 2010.
- [8] A. Majumdar, Scale-rotation invariant features from Non-Subsampled Contourlets, *Proc. of the 2015 Eighth International Conference on Advances in Pattern Recognition (ICAPR)*, IEEE, pp. 1–6, 2015.
- [9] M. Cheng, N. J. Mitra, X. Huang, et al, Global contrast based salient region detection, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 37, no. 3, pp. 569–582, 2015.
- [10] R. Achanta, S. Hemami, F. Estrada, et al, Frequency-tuned Salient Region Detection, *Proc. of the 2009 IEEE International Conference on Computer Vision and Pattern Recognition (CVPR 2009)*. IEEE, pp. 1597–1604, 2009.
- [11] James Z. Wang Research Group Database, <http://wang.ist.psu.edu/docs/> related.shtml, 2001.
- [12] Ground Truth Database, <http://www.cs.washington.edu/groundtruth/>, May 2008.
- [13] MSRA10K Salient Object Database, <http://mmcheng.net/msra10k/>, July 2014.
- [14] F. Ahmed, M. Y. Siyal, and V. U. Abbas, A secure and robust hash-based scheme for image authentication, *Signal Processing*, vol. 90, no. 5, pp. 1456–1470, 2010.
- [15] K. Fouad and J. Jianmin, Perceptual image hashing based on virtual watermark detection, *IEEE Transactions on Image Processing*, vol. 19, no. 4, pp. 981–994, 2010.
- [16] V. Monga and M. K. Mihcak, Robust and secure image hashing via non-negative matrix factorizations, *IEEE Transactions on Information Forensics and Security*, vol. 2, no. 3, pp. 376–390, 2007.
- [17] Z. Tang, S. Wang, X. Zhang, et al, Robust image hashing for tamper detection using non-negative matrix factorization, *J. of Ubiquitous Convergence and Technology*, vol. 2, no. 1, pp. 18–26, 2008.
- [18] X. Liu, Research of image perceptual hash based on scale-invariant feature transform and visual saliency, MS.D. Thesis, Chongqing: Southwest University, China, 2015.