

RESEARCH

Open Access



# Robust image hashing using ring partition-PGNMF and local features

Ram Kumar Karsh<sup>1\*</sup>, R. H. Laskar<sup>1</sup> and Bhanu Bhai Richhariya<sup>2</sup>

\*Correspondence:

ram@ece.nits.ac.in

<sup>1</sup> Department of Electronics and Communication Engineering, NIT Silchar, Silchar, Assam 788010, India  
Full list of author information is available at the end of the article

## Abstract

**Background:** Image authentication is one of the challenging research areas in the multimedia technology due to the availability of image editing tools. Image hash may be used for image authentication which should be invariant to perceptually similar image and sensitive to content changes. The challenging issue in image hashing is to design a system which simultaneously provides rotation robustness, desirable discrimination, sensitivity and localization of forged area with minimum hash length.

**Methods:** In this paper, a perceptually robust image hashing technique based on global and local features has been proposed. The Global feature was extracted using ring partition and projected gradient nonnegative matrix factorization (PGNMF). The ring partitioning technique converts a square image into a secondary image that makes the system rotation invariant. The PGNMF which is usually faster than the other NMFs has been used to reduce the dimension of the secondary image to generate the shorter hash sequence. The local features extracted from the salient regions of the image help to localize the forged region in the maliciously manipulated images. The image hashing techniques that use only global features are limited in discrimination.

**Results:** The experimental results reveal that the proposed image hashing method based on global and local features provides better discrimination capability. The proposed hashing method is tested on large image sets collected from the different standard database. It is observed from the experimental results that the proposed system is robust to content-preserving operations and is capable of localizing the counterfeit area.

**Conclusions:** The combination of global and local features is robust against the content-preserving operations, which has a desirable discriminative capability. The proposed system may be used in image authentication, forensic evidence, and image retrieval, etc.

**Keywords:** Image hashing, Multimedia security, PGNMF, Ring partition, Saliency detection

## Background

The extensive use of the Internet and multimedia services has made the image and video processing to be a part of our day-to-day activities. On the other hand, the academia and industries are facing several challenges due to the malicious manipulations of the images using image editing tools. Recognizing similar versions of the image from huge multimedia records is still a challenging issue (Wu et al. 2007). Attempt to overcome these

challenge has led to an emerging multimedia technology known as image hashing. It is a technique that extracts a short sequence from the image to represent its contents, which is used for various image processing applications (Ahmed et al. 2010; Hassan et al. 2012; Lu et al. 2004; Lu and Wu 2010; Lv and Wang 2009; Qin et al. 2012; Slaney and Casey 2008; Tang et al. 2008).

Hash function serves as an effective tool for message authentication in cryptographic applications. The hash function [such as a message digest-5 (MD-5) and a secure hash algorithm-1 (SHA-1)] defines a mapping from an arbitrary-length message to a short digest (i.e. hash string) (Stamp 2006). The prerequisite condition to accomplish the cryptographic authentication using a hashing is the avalanche effect, in which a single bit modification on the message may lead to the significant changes in the hash string. However, the over-sensitivity of the cryptographic hash function limits its applications in the multimedia domain. In the multimedia applications, the hash function should be robust against the content-preserving operations, such as the geometric transformation, format conversion, and an image compression, etc. In other words, the hash of original image and its perceptually similar version should be approximately same. Hash should be considerably different only when the visual content is distorted by the malevolent process such as object deletion/insertion operations etc. One more important property is the discriminative capability, i.e. images with dissimilar content should provide different hashes.

Various state-of-the-art image hashing methods have been proposed in the literature. Monga et al. (2006) proposed a two-stage method which consists of feature extraction and dimension reduction using clustering technique to design the final hash. It has become a commonly adopted approach for further developing the image hashing techniques. The state-of-the-art methods are either based on global (Lei et al. 2011; Swaminathan et al. 2006; Tang et al. 2008; Xiang et al. 2007) or local features (Ahmed et al. 2010; Fouad and Jianmin 2010; Khelifi and Jiang 2010; Lv and Wang 2012; Monga and Mihcak 2007; Tang et al. 2011). The image hashing techniques based on only the global feature generally provides shorter hash length, but fails to identify the local changes in the image. On the other hand, the image hashing techniques based on only the local features are generally sensitive to local changes in the image, but provides longer hash length. Xiang et al. (2007) developed an image hashing technique using image histogram that is invariant to geometric transformations. This method generally misclassifies the different images having similar histograms. Tang et al. (2008) proposed an image hashing algorithm using nonnegative matrix factorization (NMF). In this method, the secondary image is obtained by pseudo-randomly re-arranging the pixels of the normalized image. The secondary image is then passed through the NMF, the computed coefficients are then quantized and scrambled to generate the final hash. Swaminathan et al. (2006) developed a Fourier-Mellin transform based method which is rotation invariant and provides better security. Lei et al. (2011) proposed an image hashing technique based on discrete Fourier transform (DFT) and radon transform (RT). DFT is carried out on the invariant moments generated by the Radon transformation of the original image. The computed DFT coefficients are then quantized to form the final hash. Khelifi and Jiang (2010) calculated image hash on the basis of the virtual watermark detection, which is robust for the geometric deformations and content preserving operations. This method

can detect only the changes in large areas. An NMF–NMF based technique has also been proposed to generate image hash (Monga and Mihcak 2007). First NMF provides a secondary image from pseudo-randomly selected sub-images and the second NMF is used to generate the low-rank matrix. Then the matrix entries are concatenated to form the NMF–NMF vector which is considered to be the final hash. This method fails to identify the counterfeit areas. An analysis of NMF–NMF method suggests that out of the three keys, the first one used for pseudo-randomly selecting sub-images is only important (Fouad and Jianmin 2010).

Tang et al. (2011) proposed a lexicographical-framework for image hashing that uses discrete cosine transform (DCT) and NMF. This method is robust to some content preserving operations, but fragile to large degree rotation. Ahmed et al. (2010) has developed an image hashing using discrete wavelet transform (DWT) and SHA-I. It has been used for tamper detection, but limited in brightness/contrast changes and rotation. An image hashing technique based on scale invariant feature transform (SIFT) and Harris detector has also been proposed Lv and Wang (2012). In this method, key points are extracted from the scale space image by using SIFT and from which the stable key points are retained by Harris detector. The shape-texture features around the stable key points are used to generate the final hash. It is robust to the geometric deformations and can also find the counterfeit area. This method does not possess rotation invariant property.

Lu et al. (2010) developed a concept of forensic hash for information assurance. It is based on RT and the scale space theory. For image authentication, it uses the side information. This method can provide the history of image manipulations and parameters for geometric deformation. It can also identify the counterfeit area. An image hashing method with the combination of global and local features has been proposed in the literature (Zhao et al. 2013). These features are based on Zernike moments and shape-texture respectively. It is robust to the most content preserving operations and can also locate the forged areas. But it is not robust to the large degree of rotation.

Tang et al. (2016a, b) proposed an image hashing based on ring partition and invariant vector distance by incorporating invariant vector distance in place of NMF. This method is robust against perceptual digital image manipulation, but it cannot localize the forged region. The entropy based method (Tang et al. 2013) creates a rotation invariant secondary image based on the ring partition and extracts entropy from each ring to generate the image hash. This method is robust to digital image manipulations and possesses good discrimination capability for the natural image, but limited discrimination capability for textural images.

Sun and Zeng (2014) designed an image hashing based on combining the Fourier-Mellin transform (FMT) and the compressive sensing (CS). This method is resilient to scaling, and translation attacks, but sensitive to large angle rotation. The global features extracted from partitioned image blocks and local feature extracted from key points of the image have been concatenated to form an image hash (Wang et al. 2015). It is robust to perceptual image manipulation, can locate the forged area but sensitive to large angle rotation. An image hashing method based on the quaternion discrete Fourier transform followed by log-polar transform has been developed (Ouyang et al. 2015). In this method, the three colors have been used to generate hash without increasing the hash length. This method is incapable of locating the counterfeit areas in the forged image.

Yan et al. (2016) proposed an image hashing method using adaptive local features. It is robust against the content-preserving operations, but unable to detect color forgery. Qin et al. (2016) has designed a hashing method based on the block truncation coding (BTC). The major drawback of this method is sensitive towards brightness changes. An image hashing technique based on ring partition and color vector angle has been proposed by Tang et al. (2016a, b). This method is robust to digital image manipulation, but around 20% of the corner information has not been utilized to design the image hash.

Image rotation is a normal content preserving operation that is considered to be a major challenging issue in the case of image hashing. Most of the conventional image hashing methods categorize the rotated images as different images (Fridrich and Goljan 2000; Qin et al. 2012; Lin and Chang 2001; Tang et al. 2011; Venkatesan et al. 2000; Wu et al. 2009). There are some existing methods that are resilient to the rotation but provides the poor discriminative capability (Kozat et al. 2004; Lefebvre et al. 2002; Lei et al. 2011; Monga and Mihcak 2007; Roover et al. 2005; Swaminathan et al. 2006). Even though the ring partition and NMF based method (Tang et al. 2014) having good discrimination capability, is invariant to arbitrary rotation, yet it cannot localize the counterfeit region of the forged image. Localization of the counterfeit regions in the forged image is very useful in the medical and forensic evidence (Mishra and Adhikary 2013). This method also suffers from the convergence problem of the matrix factorization, which is based on the multiplicative update method. It is an extremely challenging task to design an image hashing technique that simultaneously satisfies all the requirements, such as rotation robustness, good discriminative capability, and localization of the forged area.

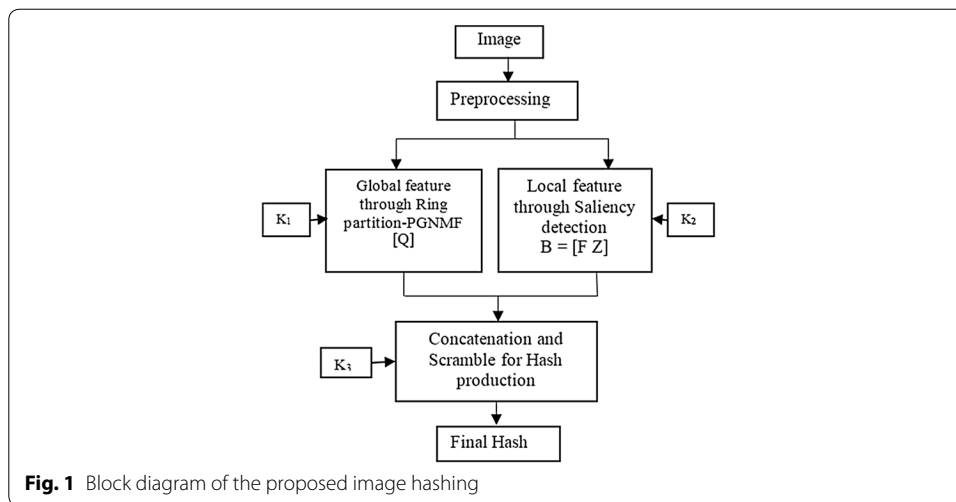
In this paper, a robust image hashing technique has been proposed to meet the above requirements simultaneously. In the proposed method, the secondary image obtained through ring portioning makes the system rotation invariant. The use of PGNMF instead of NMF helps the system to converge faster. The attribute of the PGNMF (i.e. learning parts-based representation) and use of the local features along with the global features helps to enhance the discrimination capability of the proposed algorithm. The combination of local features with the global features also makes the system capable of identifying the forged regions in the manipulated images.

Experiments were carried out on large image sets that were taken from standard image databases (USC-SIPI Image Database 2007; Ground Truth Database 2008) and downloaded from the internet. The experimental results show that the proposed algorithm provides better tradeoff amongst the rotation robustness, discrimination capability and localizing the counterfeit region.

## Methods

The proposed image hashing involves four different steps as shown in Fig. 1.

The input image is transformed to a normalized image size ( $256 \times 256$ ) with pre-processing. The final hash is formed by concatenating the two intermediate hash. First intermediate hash, is generated from the global features derived from the normalized image using ring partitioning and PGNMF. Second intermediate hash is generated from the shape-texture feature derived from the salient regions of the image (Hou and Zhang 2007). The final hash has been scrambled using the secret key  $K_3$ . The PGNMF, ring



partition, saliency detection scheme, local features and the proposed algorithm in details with a probable application have been discussed in the subsequent subsections.

**PGNMF**

Non-negative matrix factorization is an algorithm where a matrix is factorized into two matrices of lower dimension, with the property that all matrices have only positive elements (Paatero and Tapper 1994). Let  $L$  be a matrix of size  $m \times n$ , where  $L_{ij} \geq 0$ , then it can be approximately represented by the two non-negative matrices  $W \in \mathbb{R}^{m \times p}$  (i.e. base matrix) and  $H \in \mathbb{R}^{p \times n}$  (i.e. coefficient matrix) by using NMF such as:

$$L \approx WH \tag{1}$$

where  $p$  is the rank of NMF that should satisfy the condition  $p < \min(m, n)$ .

If each column of the matrix  $L$  represents the objects of an image, then NMF approximate it by linear combination of  $p$  basis columns in  $W$ . There are various applications of NMF such as document clustering (Xu et al. 2003), finding basis vectors of the images (Lee and Seung 2001), and molecular pattern discovery (Brunet et al. 2003). The theoretical issues related to NMF scheme have been discussed in Donoho and Stodden (2004). The state-of-the-art NMF schemes have the convergence problem. To overcome this problem, Lin (2007) proposed a method, named PGNMF which has the faster optimization property. The illustration of PGNMF scheme is shown in Table 1.

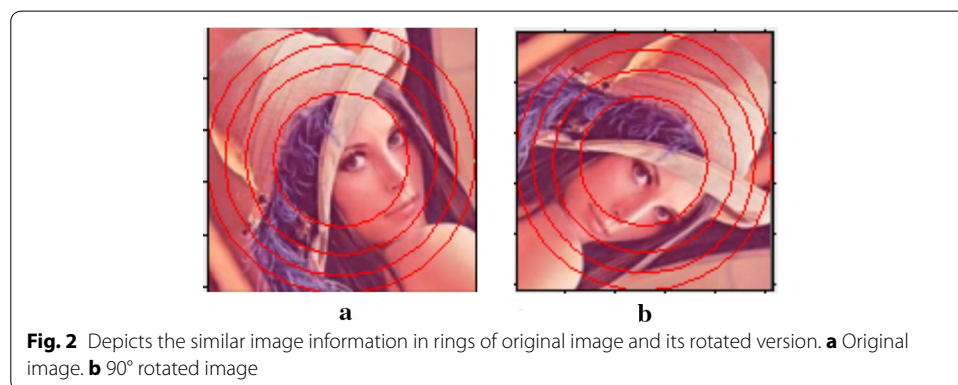
**Square image to secondary image (circular) conversion**

Generally, image center remains fix after rotation manipulation. It has been observed that pixel values in different rings of the original image and its rotated version remain same as shown in Fig. 2a, b respectively. Hence, the square image can be converted into rings that may be subsequently used to construct the secondary image as shown in Fig. 3. Use of this concept makes the proposed system rotation invariant (Tang et al. 2014). The secondary image is constructed as follows:

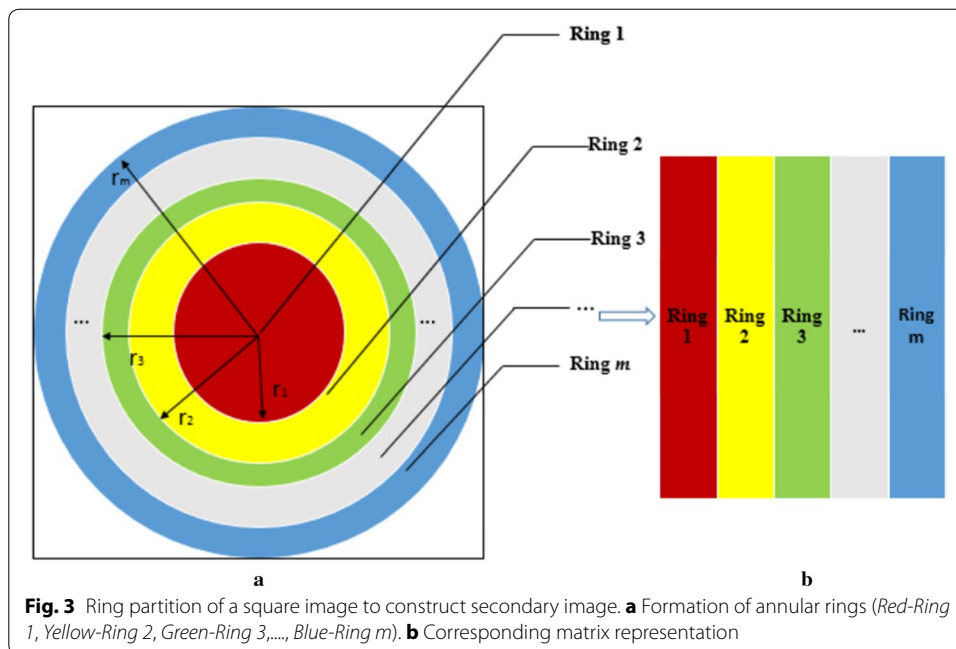
Let the size of a square image be  $n \times n$ , the number of rings is  $m$ , and  $P_i$  be the set of pixel values in the  $i_{th}$  ring ( $i = 1, 2, \dots, m$ ). The inscribed circle area from a square image

**Table 1 The PGNMF algorithm**

<p><i>Input:</i> A non-negative matrix <math>\mathbf{L}</math></p> <p><i>Output:</i> Two non-negative matrix <math>\mathbf{W}</math> and <math>\mathbf{H}</math></p> <p>The approach is to find <math>W</math> and <math>H</math> by minimizing the difference between <math>\mathbf{L}</math> and <math>\mathbf{WH}</math>:</p> $\min_H f(\mathbf{H}) = \frac{1}{2} \ \mathbf{L} - \mathbf{WH}\ _F^2 \quad (2)$ <p>Subject to <math>H_{bj} \geq 0, \forall b, j</math>, where <math>\ \cdot\ _F</math> is Frobenius norm</p> <p>Here, <math>\mathbf{L}</math> and <math>\mathbf{W}</math> are constant matrices.</p> <p>We can solve (2) using an algorithm: an improved projected gradient method</p> <ol style="list-style-type: none"> <li>1. Given <math>0 &lt; \beta &lt; 1; 0 &lt; \sigma &lt; 1</math>. Initialize any feasible <math>\mathbf{H}^1, \beta</math>, set <math>\alpha_0 = 1</math>.</li> <li>2. for <math>k = 1, 2, \dots</math> <ol style="list-style-type: none"> <li>(a) Assign <math>\alpha_k \leftarrow \alpha_{k-1}</math></li> <li>(b) If <math>\alpha_k</math> satisfies (3)                     <math display="block">f(\mathbf{H}^{k+1}) - f(\mathbf{H}^k) \leq \sigma \nabla f(\mathbf{H}^k)^T (\mathbf{H}^{k+1} - \mathbf{H}^k) \quad (3)</math> <p>repeatedly increase <math>\alpha_k</math> by</p> <math display="block">\alpha_k \leftarrow \alpha_k / \beta</math> <p>until either <math>\alpha_k</math> does not satisfy (3) or <math>\mathbf{H}(\alpha_k / \beta) = \mathbf{H}(\alpha_k)</math></p> <p>else repeatedly decrease <math>\alpha_k</math> by</p> <math display="block">\alpha_k \leftarrow \alpha_k \cdot \beta</math> <p>until <math>\alpha_k</math> satisfy (3)</p> </li> <li>(c) Set <math>\mathbf{H}^{k+1} = P[\mathbf{H}^k - \alpha_k \nabla f(\mathbf{H}^k)]</math></li> </ol> </li> </ol> <p>Where <math>P[x_i] = \begin{cases} x_i &amp; \text{if } l_i &lt; x_i &lt; u_i \\ u_i &amp; \text{if } x_i \geq u_i \\ l_i &amp; \text{if } x_i \leq l_i \end{cases}</math></p> <p>The same procedure can be used to obtain <math>\mathbf{W}^{k+1}</math></p> $\min_W f(\mathbf{W}) \equiv \frac{1}{2} \ \mathbf{L}^T - \mathbf{W}^T \mathbf{H}^T\ _F^2 \quad (4)$ <p>Subject to <math>W_{bj} \geq 0, \forall b, j</math></p> <p>where <math>\mathbf{L}^T</math> and <math>\mathbf{H}^T</math> are constant matrices.</p>
---



has been used for hash generation. The inscribed area has been divided into annular rings of equal area as shown in Fig. 3a. Each ring provides the column of the secondary image as shown in Fig. 3b. The radii of circles can be determined as follows: For an image



of size  $n \times n$ , the outermost radius is  $r_m = \lfloor \frac{n}{2} \rfloor$ . To find the other radii, firstly the area  $S$  enclosed by outermost radius is calculated and then the other ring area  $U_A$ , i.e. average area is found as follows (Tang et al. 2014):

$$S = \pi r_m^2 \tag{5}$$

$$U_A = [S/m] \tag{6}$$

So,  $r_1$  can be computed by

$$r_1 = \sqrt{\frac{U_A}{\pi}} \tag{7}$$

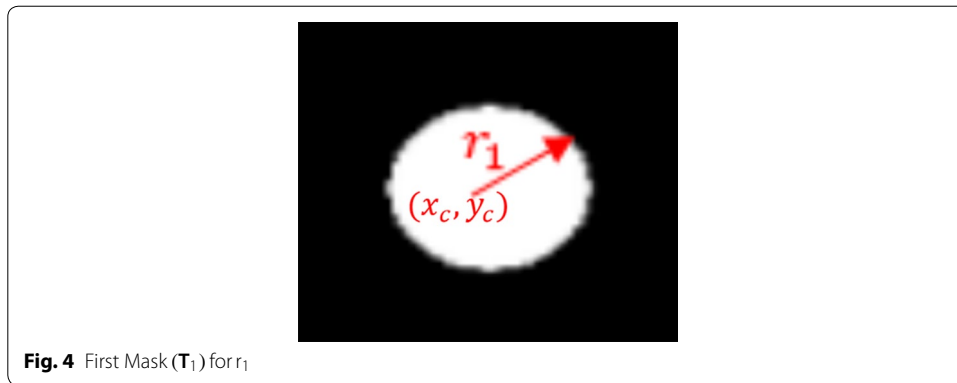
Thus, other radii  $r_k (k = 1, 2, \dots, m - 1)$ , can be found using the following equation:

$$r_k = \sqrt{\frac{U_A + \pi r_{k-1}^2}{\pi}} \tag{8}$$

The process of arranging the pixels from the rings (shown in Fig. 3a) into columns (shown in Fig. 3b) is as follows: Let the coordinates of the image center be  $(x_c, y_c)$ . If  $n$  is an even number, then,  $x_c = \frac{n}{2} + 0.5$  and  $y_c = \frac{n}{2} + 0.5$ , otherwise,  $x_c = \frac{(n+1)}{2}$  and  $y_c = \frac{(n+1)}{2}$ .

A circular region ( $C_1$ ) (i.e. Ring 1) of radius  $r_1$  around the center pixel has been selected as shown in (9)

$$C_1 : (x - x_c)^2 + (y - y_c)^2 = r_1^2 \tag{9}$$



where  $y$ th row and the  $x$ th column of the normalized image is given by  $1 \leq x \leq n, 1 \leq y \leq n$ .

The process of creating a mask ( $T_1$ ) for the first radius  $r_1$  is given below:

First mask ( $T_1$ ) is constructed having pixel value one (*true*) inside the ring of radius  $r_1$  and pixel value zero (*false*) outside the circle as shown in Fig. 4.

Let the pixel values of this region (i.e. Ring 1) be placed in the first column of matrix  $L$  by multiplying the input image ( $I$ ) and first mask ( $T_1$ ):

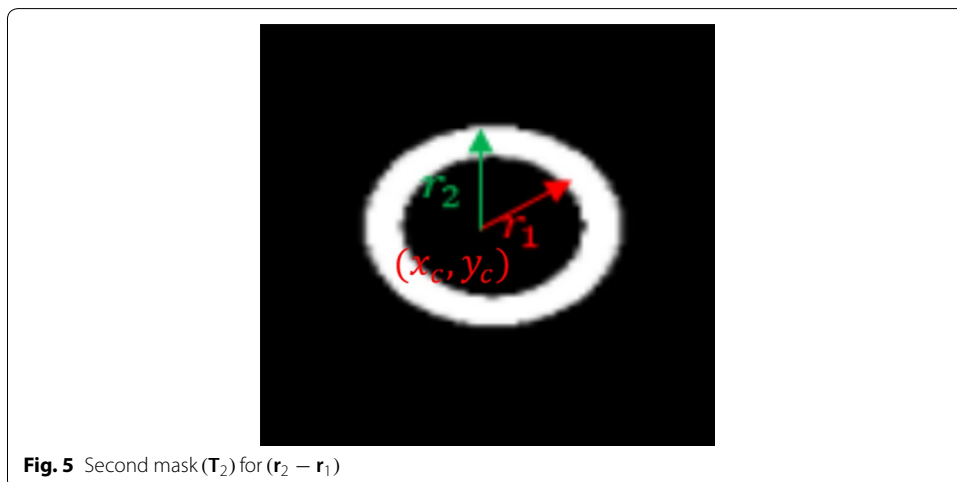
$$L_1 = I \times T_1; \quad \text{inside } r_1 \tag{10}$$

Similarly, other circles  $C_k (k = 2, 3, \dots, m)$  can be constructed for  $r_k (k = 2, 3, \dots, m)$ .

$$C_k : (x - x_c)^2 + (y - y_c)^2 = r_k^2 \tag{11}$$

The second mask ( $T_2$ ) is constructed with pixel value one (*true*) for the pixels inside the annular ring of  $r_1 \leq r \leq r_2$  and pixel value zero (*false*) outside the annular ring as shown in Fig. 5.

The pixel values of this region is placed (i.e. annular Ring 2) into the second column of matrix  $L$  by multiplying the input image ( $I$ ) and second mask ( $T_2$ ):





$$\mathbf{L}_2 = \mathbf{I} \times \mathbf{T}_2; \quad \text{for } (\mathbf{r}_2 - \mathbf{r}_1) \tag{12}$$

Similarly, pixel values of other regions (i.e. annular Rings) are putted in the other column of matrix  $\mathbf{L}$ :

$$\mathbf{L}_k = \mathbf{I} \times \mathbf{T}_k; \quad \text{for } (\mathbf{r}_k - \mathbf{r}_{k-1}) \tag{13}$$

Thus, the rotation invariant matrix is obtained as follows:

$$\mathbf{L} = [\mathbf{L}_1, \mathbf{L}_2, \mathbf{L}_3, \dots, \mathbf{L}_m] \tag{14}$$

To compress a high dimensional vector, the matrix  $\mathbf{L}$  is processed through PGNMF to obtain a compact form of first intermediate hash ( $\mathbf{Q}$ ).

**Saliency map recognition**

The salient region is defined as the visual attention of any image. For every image, there are two parts, an innovation and a prior knowledge (Hou and Zhang 2007). It is also called the novel and redundant parts. Redundant parts should be removed to get the novel parts of an image.

The log-spectrum,  $V(f)$  characterize the common information of an image. The redundant information is present in  $V(f)$  due to the fact that the log-spectra of different images are similar. The convolution of  $V(f)$  and an  $l \times l$  low-pass kernel is used to find the redundant information  $R(f)$ , as given by (15)

$$R(f) = h_l \times V(f) \tag{15}$$

The novel portion ( $N(f)$ ) of an image can be obtained by spectral residual method that can be derived by subtracting  $R(f)$  from  $V(f)$  and the salient map can be obtained by inverse Fourier transform of  $f$ :

$$S_m(y) = F^{-1}[N(f)] = F^{-1}[V(f) - R(f)] \tag{16}$$

The salient region of an image can be obtained by taking the threshold three times of the mean of the  $S_m(y)$  (Zhao et al. 2013).

**Local features**

For human visual perception, the important local feature is the texture of the image. Six texture features have been proposed in Deselaers et al. (2004), Tamura et al. (1978). We have used three features, namely, coarseness ( $\delta_1$ ), contrast ( $\delta_2$ ), and kurtosis ( $\delta_3$ ). To calculate coarseness with respect to a particular pixel( $i, j$ ), the average of its neighborhood size ( $2^k \times 2^k$ ) is considered as follows:

$$A_k(i, j) = \frac{1}{2^{2k}} \sum_{x=i-2^k}^{i+2^k-1} \sum_{y=j-2^k}^{j+2^k-1} g(x, y), \quad k = 1, 2, \dots, 5 \tag{17}$$

Here,  $g(x, y)$  represents the gray-level pixel. At every pixel( $i, j$ ), differences of average values in vertical and horizontal direction is as follows:

$$E_{k,h}(i, j) = \left| A_k(i + 2^{k-1}, j) - A_k(i - 2^{k-1}, j) \right| \tag{18}$$

$$E_{k,v}(i, j) = \left| A_k(i, j + 2^{k-1}) - A_k(i, j - 2^{k-1}) \right| \tag{19}$$

The  $S_{best}(i, j)$  represents the highest difference value evaluated at  $(i, j)$ , and is given by (20)

$$S_{best}(i, j) = \arg \max_{k=0, \dots, 5; d=h, v} E_{k,d}(i, j) \tag{20}$$

The coarseness ( $\delta_1$ ) for the image can be calculated as

$$\delta_1 = \frac{1}{m \times n} \sum_i^m \sum_j^n S_{best}(i, j) \tag{21}$$

where,  $m$  and  $n$  represents the dimension of the image.

Image illumination changes are described as contrast. This can be evaluated using the variance ( $\sigma^2$ ) and the fourth order moment ( $\mu_4$ ) of the gray value for the given region.

$$\delta_2 = \sigma^2 \mu_4^{-1/4} \tag{22}$$

The kurtosis which is used to calculate polarization of image can be measured as

$$\delta_3 = \frac{\mu_4}{\sigma^4} \tag{23}$$

**Brief description of the proposed image hashing**

The detail of the proposed image hashing has been illustrated in Table 2.

The length of the final hash has been discussed in Table 3. It is observed that the final hash length is 106 integers. As the dynamic range of integer is [0, 255], therefore, it requires 8 bit to represent each integer value. Hence, the final length of the hash in bits is  $106 \times 8 = 848$  bits.

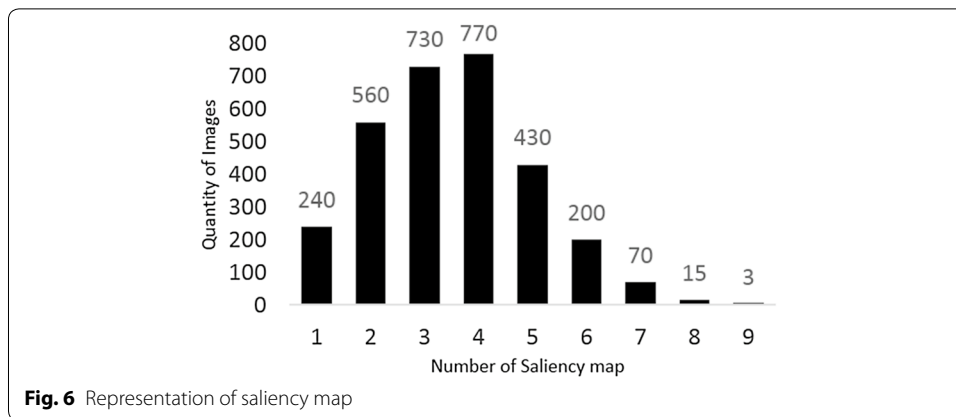
**Authentication of images**

For an image authentication, we have the reference hash ( $\mathbf{M}_0$ ) of the original (reference) image, calculated hash ( $\mathbf{M}_1$ ) for the received (test) image. The above two hashes are compared to check whether the received image is similar/tampered version of the original image or simply the different one. A pair of visually similar images may have different pixel values due to the content-preserving operations. But visually same are called similar or same image. The process of image authentication is as follows:

1. *Extracting the features* The final hash for the test image is calculated using the proposed algorithm as discussed in Table 2 and let the hash be named as  $\mathbf{M}_1 = [\mathbf{Q}_1 \mathbf{F}_1 \mathbf{Z}_1]$ .
2. *Separation of final hash* The reference hash ( $\mathbf{M}_0$ ) obtained from the original image is decomposed into the global and local feature vectors to obtain the components  $\mathbf{Q}_0, \mathbf{F}_0$ , and  $\mathbf{Z}_0$
3. *Saliency map comparison* The salient regions of the test image are compared with those of the original image using position vectors. If it is observed that the compared

**Table 2 Proposed image hashing in details (From input image to generation of final hash)**

<p><i>Input: Image</i></p> <ol style="list-style-type: none"> <li>1. Pre-processing: The input image <math>\mathbf{I}</math> is rescaled to a fixed size <math>n \times n</math> by bilinear interpolation, and changed from the RGB to the YCbCr representation. The Y and <math> \text{Cb} - \text{Cr} </math> known as luminance and chrominance constituents of the image which may be used to create the hash. However, we have selected <math>n = 256</math> and Y channel only for this study</li> <li>2. Global features extraction using ring partition-PGNMF             <ol style="list-style-type: none"> <li>a. Let us consider the Y component of the pre-processed image;</li> <li>b. Divide Y into <math>m</math> rings and produce a secondary image <math>\mathbf{L}</math> by using Eqs. (5–14)</li> <li>c. Apply PGNMF to <math>\mathbf{L}</math>, and obtain the coefficient matrix <math>\mathbf{H}</math></li> <li>d. Concatenate the matrix coefficients to give the hash of length <math>\mathbf{Q}' = nk</math> by global feature</li> <li>e. A secret key <math>K_1</math> randomly generates a row vector <math>\mathbf{X}_1</math>; The first intermediate hash <math>\mathbf{Q}</math> is formulated as <math>\mathbf{Q} = [(\mathbf{Q}' + \mathbf{X}_1) \bmod 256]</math></li> </ol> </li> <li>3. Extraction of local features             <ol style="list-style-type: none"> <li>a. Detect <math>C</math> numbers of largest salient regions from pre-processed image <math>Y</math></li> <li>b. The <math>C</math> numbers of position vectors of four dimensions are formed using the coordinates of the top left corner, and width/height of each rectangle around the salient regions. It is denoted by position vector <math>\mathbf{f}^k</math> (<math>k = 1, 2, \dots, C</math>)</li> <li>c. The texture features are computed for <math>C</math> number of salient regions. Fig. 6 shows that 96% of the images have less than 6 salient regions. We have used 2900 images to test the number of salient regions as shown in Fig. 6. Only a few of the images have more than 6 salient regions. There is a trade-off between the hash length and number of salient regions. When the value of <math>C</math> is large, fewer salient regions will be left out but will produce a longer image hash. We have taken <math>C = 6</math>, empirically as an optimal trade-off between hash length and the salient maps that will be left out. For every salient region, local texture features, that include coarseness (<math>\delta_1</math>), contrast (<math>\delta_2</math>) and kurtosis (<math>\delta_3</math>) are computed and rounded to form <math>C</math> element texture vectors <math>\mathbf{z}^k</math> (<math>k = 1, 2, \dots, C</math>) of three dimensions</li> <li>d. The position vector <math>\mathbf{F}</math> is concatenated with the texture vector <math>\mathbf{Z}</math> to give local feature vector <math>\mathbf{B}' = [\mathbf{f} \dots \mathbf{f}^C \dots \mathbf{z}^1 \dots \mathbf{z}^C]</math>. For an image having lesser than 6 salient regions, the position and texture features are set to zero for the missing ones</li> <li>e. A secret key <math>K_2</math> randomly generates a row vector <math>\mathbf{X}_2</math>. The second intermediate hash <math>\mathbf{B}</math> is formulated as <math>\mathbf{B} = [(\mathbf{B}' + \mathbf{X}_2) \bmod 256]</math></li> </ol> </li> <li>4. Finally, the two intermediate hash vectors are concatenated and pseudo-randomly scrambled on the basis of secret key <math>K_3</math> to obtain the final hash <math>\mathbf{M}</math></li> </ol> <p><i>Output: Final Hash</i></p>
--



**Table 3** Structure of image hash

Global vector	Salient vector S		
Ring partition-PGNMF	$F(x, y, w, h)$	Z	Final hash length
64 integers	$4 \times 6 = 24$ integers	$3 \times 6 = 18$ integers	106 integers

areas for a pair of salient regions are large enough, then it may be considered to be matched. Accordingly, the texture vectors are reshuffled by moving the matched components in each of the texture vector pairs  $Z_0, Z_1$ .

Consider an example where, there are five salient regions in the reference image and four in the test image.

$$Z_0 = [z_0^1 z_0^2 z_0^3 z_0^4 z_0^5 0]$$

$$Z_1 = [z_1^1 z_1^2 z_1^3 z_1^4 0 0]$$

The first four pair of sub-vectors in  $Z_0$  and  $Z_1$  may be either match or unmatched. Accordingly,  $F_0$  and  $F_1$  are reshuffled.

4. *Calculating Euclidean distance and decision making* Euclidean distance between the image hashes is used to judge resemblance of the image pairs. Let us consider an image hash  $N$ , which is structured by concatenating the global vector  $Q$  and the rearranged texture feature  $Z$ , say  $N = [Q, Z]$ . Position feature ( $F$ ) is not required for distance calculation. However, it will be used during the counterfeit area localization. The Euclidean distance between reference hash vector ( $N_0$ ) and the test hash vector ( $N_1$ ) is defined as:

$$D = \|N_1 - N_0\| \tag{24}$$

The Euclidean distance is used to decide the similarities or dissimilarities with respect to the pre-determined threshold. In practice, the global feature of an image obtained by applying Ring partition-PGNMF approach is adequate enough to distinguish the similarities/dissimilarities of the image pairs. The inaccuracy of the saliency detection may have an undesirable influence on the calculation of Euclidian distance that may lead to

erroneous results. To overcome this problem, the hash distance may be calculated based on global features only to detect the similarities/dissimilarities in the decision-making process. This is given as follows.

$$D = \|Q_1 - Q_0\| \tag{25}$$

**Results**

The experimental results for the robustness, discriminative capability, and counterfeit area localization have been discussed in the next subsections.

**Robustness**

Robustness evaluations are quantitative analysis for the normal content-preserving operations. An image hash should be approximately same after any non-malicious activity. These activities consist of image rotation, JPEG compression, watermark embedding, Gaussian low-pass filtering, gamma correction, brightness adjustment, contrast adjustment, and image scaling (Swaminathan et al. 2006).

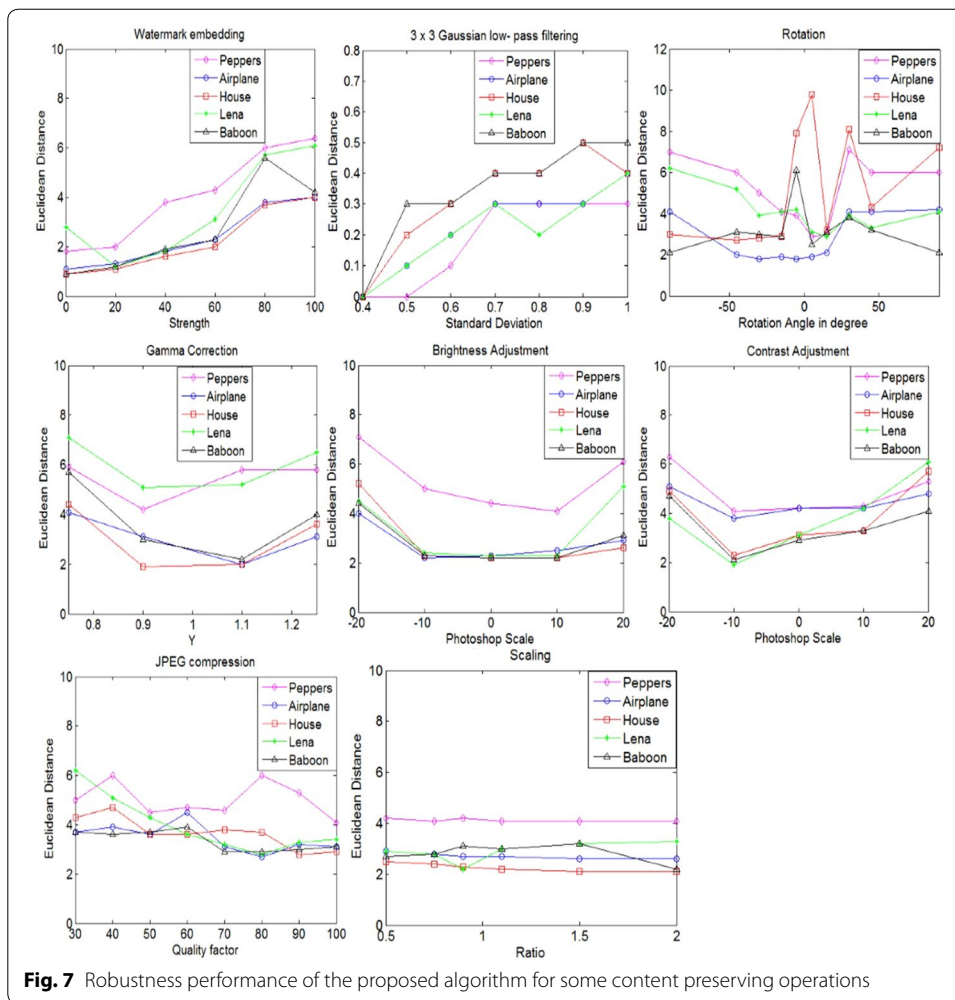
We have tested the proposed model using 200 different images taken from various sources (USC-SIPI Image Database 2007; Ground Truth Database 2008). Because of space limitation, we have used only five images (Peppers, Airplane, House, Lena, and Baboon) for visual representation in the paper. Content preserving operation is created by StirMark 4.0 (Petitcolas 2000), Photoshop CS3 (10.0), and MATLAB (2013a) as shown in Table 4.

The Euclidean distance between the original image and the image after content preserving operations is shown in Fig. 7. It may be observed that if we choose the threshold of Euclidean distance as 7, then few perceptually similar image (i.e.  $3/200 \times 100\% = 1.5\%$ ) may be considered as the different image, as shown in Fig. 7 (rotation operation). Therefore, the true rejection is 1.5%. However, the false acceptance is zero percent.

The Euclidean hash distances for visually different images are almost greater than 10 as shown in Fig. 9. If we increase the selected threshold of the Euclidean distance to 10, the true rejection is zero percent. Nevertheless, there are few cases (0.116%) in which the Euclidean distance is below 10 for different images. Thus, the false acceptance is 0.116%. Hence, it may be concluded that the threshold of the Euclidean distance may be selected according to the application of an image hashing.

**Table 4 Content-preserving operations with given parameter values**

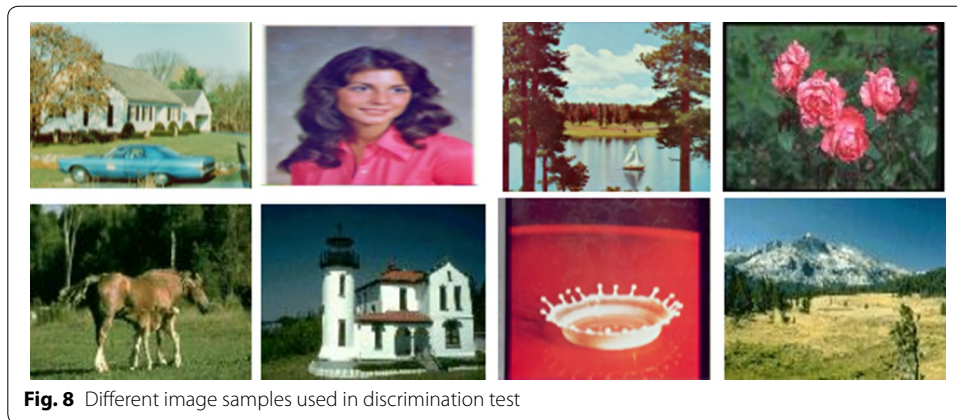
Tool	Manipulation	Parameter	Parameter values
Stir Mark	JPEG compression	Quality factor	30, 40, 50, 60, 70, 80, 90, 100
Stir Mark	Watermark embedding	Strength	10, 20, 30, 40, 50, 60, 70, 80, 90, 100
Stir Mark	Scaling	Ratio	0.5, 0.75, 0.9, 1.1, 1.5, 2.0
Stir Mark	Rotation and Cropping	Rotation angle in degree	1, 2, 5, 10, 15, 30, 45, 90, -1, -2, -5, -10, -15, -30, -45, -90
Photoshop	Brightness adjustment	Photoshop Scale	10, 20, -10, -20
Photoshop	Contrast Adjustment	Photoshop Scale	10, 20, -10, -20
Matlab	Gamma Correction	Y	0.5, 0.75, 0.9, 1.1, 1.5
Matlab	3 × 3 Gaussian low pass filter	Standard deviation	0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1



For this study, we have used an image hashing technique for an image authentication purpose. The basic requirement of an image authentication is to have zero false acceptance rate. Because of this, we have selected a threshold Euclidian distance 7 for subsequent analysis. The experimental results which are shown in Fig. 7 demonstrates the effectiveness of the proposed method. It can be observed from Fig. 7, the Euclidean distances of some content preserving operations are less than the selected threshold (i.e. 7) except few cases in rotation operation. However, these few cases in rotation operation may be considered similar image at selected threshold 10, but with the cost of small false acceptance. Hence, the proposed method is robust to some content preserving operations.

### Discrimination capability

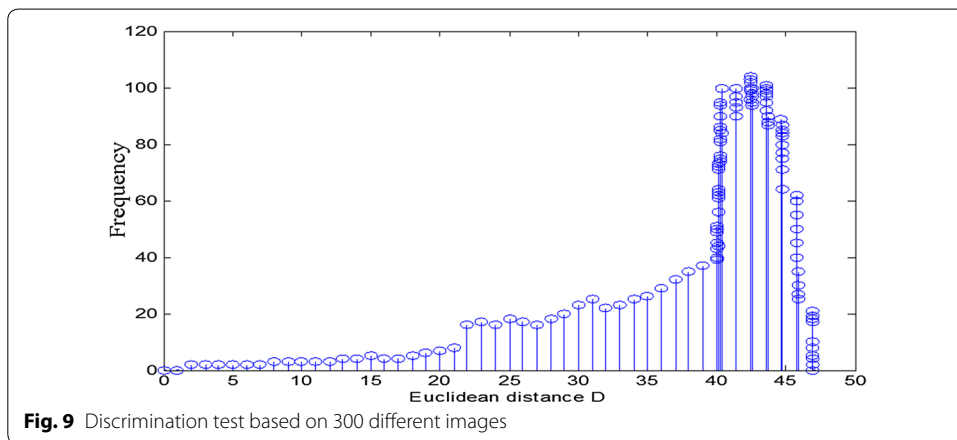
To check the discriminative capability of the proposed method for visually different images, we have taken 300 images from different sources. Some of the sample images (animal, car, house, scene and lady, etc.) from the 300 image set have been shown in Fig. 8. The hash vector for the 300 images has been generated using the proposed method and the calculated Euclidean distance for different possible combinations of

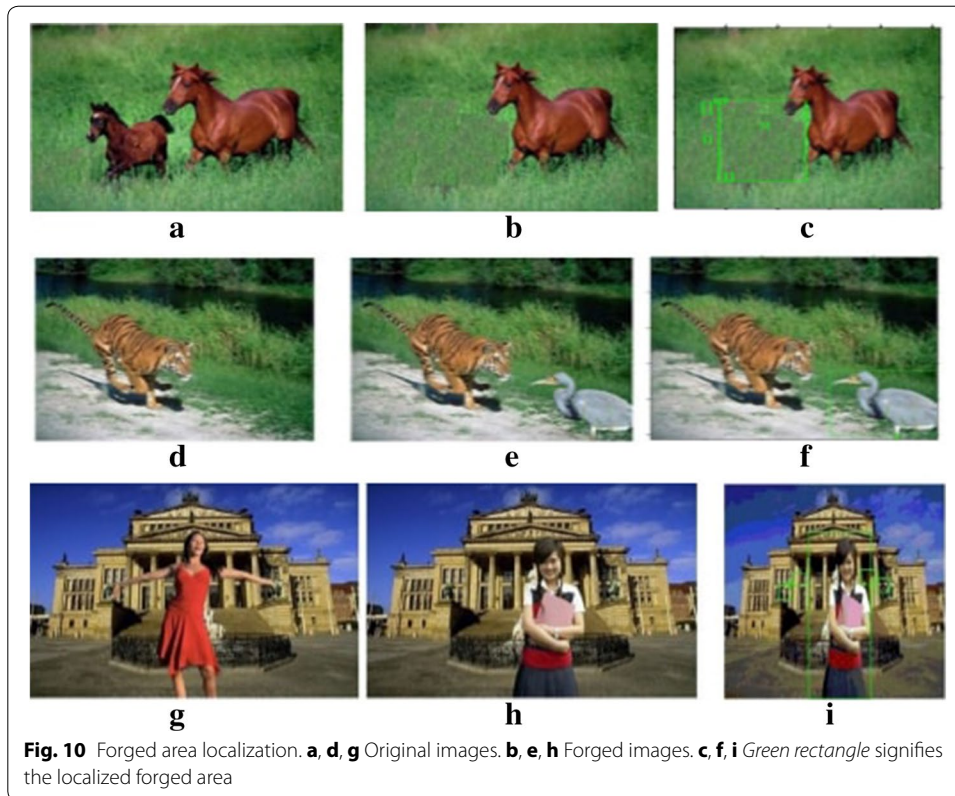


hash vector pairs are shown in Fig. 9. It can be observed that the obtained Euclidean distances are above 7 (i.e. pre-determined threshold) in all the cases. This signifies that the proposed method provides satisfactory discrimination.

**Testing for counterfeit area localization**

In addition, we have carried out the experiments for malicious activities and localization of the counterfeit area. Whenever a particular image is attacked by the malicious operation, such as an insertion, deletion, and replacement, etc., it looks perceptually different (Swaminathan et al. 2006). Then, the obtained image is called a forged image and should provide different hash values. The forged images are generated using Photoshop. Some examples of the original and forged images are shown in Fig. 10. The experimental results of forged area localization have been shown in the third column of Fig. 10. In Fig. 10c, the green rectangle is the forged area, in which some content that was there in the original image has been removed in forged image. Similarly, green rectangle in Fig. 10f, i shows the forged area localization for content insertion and object replacement respectively. We have generated the hash vectors of original images and forged images, and calculated the hash distances which are shown in Table 5. It is observed that all the





**Table 5 Image hashes for tampered image**

Image pairs	D
Figure 10a, b	31
Figure 10d, e	29
Figure 10g, h	34

hash distance are more than the pre-selected threshold. Hence, it may be concluded that the proposed method is sensitive towards malicious changes made in the image.

**Discussion**

The proposed algorithm has been compared with some of the well-known methods available in the literature (Ahmed et al. 2010; Khelifi and Jiang 2010; Monga and Mihcak 2007; Tang et al. 2008; Zhao et al. 2013). The performances have been evaluated using Receiver Operating Characteristics (ROC). The ROC is a characteristic curve plotted using false positive rate (FPR) along the abscissa and true positive rate (TPR) along the ordinate. The FPR and TPR are calculated as follows:

$$FPR = \frac{\epsilon_1}{O_1} \tag{26}$$

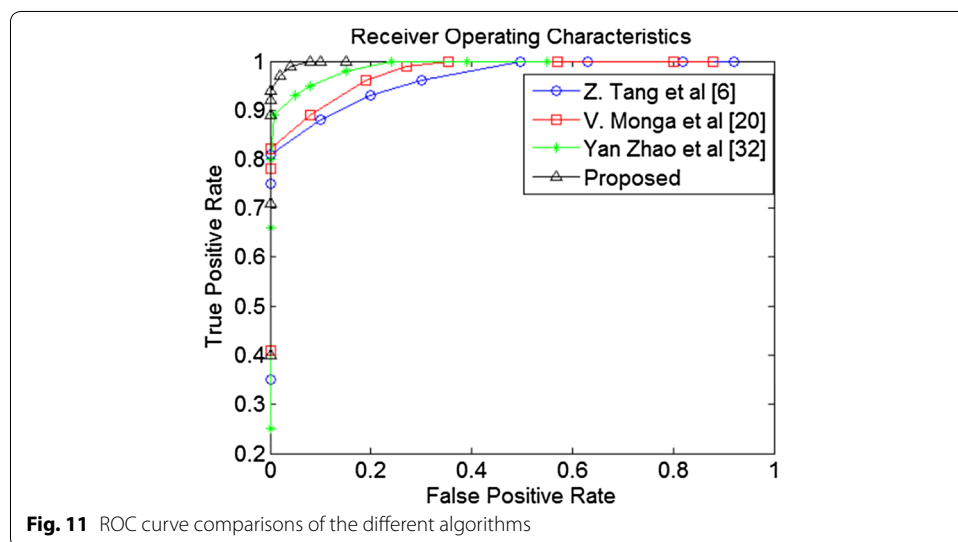


$$TPR = \frac{\epsilon_2}{O_2} \tag{27}$$

where,  $\epsilon_1$  represents the number of visually different image pairs judged as similar image pairs and  $\epsilon_2$  represents the number of perceptually identical image pairs judged as similar image pairs. The  $O_1$  and  $O_2$  denote the total number of visually dissimilar image pairs and a total number of perceptually similar image pairs respectively. The ROC comparison amongst the different algorithms is shown in Fig. 11. It can be observed from Fig. 11 that the proposed method TPR is higher especially at low FPR as compared to the other methods.

The performance comparison of the proposed method with some of the state-of-the-art methods are shown in Table 6. The methods proposed in Ahmed et al. (2010), Monga and Mihcak (2007), Tang et al. (2008) are not robust to the rotation operation. This may be due to the block-based operations carried out for feature extraction. The hashing methods proposed in Tang et al. (2008) and Monga and Mihcak (2007) fails to localize the forged area in the image. The method proposed in Khelifi and Jiang (2010) is robust to content-preserving operations but fails to find the small counterfeit area. The method proposed in Zhao et al. (2013) is robust against almost all the geometric operations but not robust to rotation operation beyond 5°. For ROC comparison, we have used only the baseline methods (used in this study) as described in Monga and Mihcak (2007), Tang et al. (2008), Zhao et al. (2013).

It may be observed from Table 6 that the proposed algorithm is robust against the image manipulation and possesses good discriminating characteristics. Moreover, it can locate the counterfeit area with sufficient accuracy. Table 6 shows that the proposed method TPR (i.e. 0.9811 with zero FPR) and FPR (i.e. 0.0012 with TPR one) is higher and lower respectively as compared to other methods. The limitation of the proposed algorithm is that full image information has not been used for hash generation. Around 21% of the image information remain unused during the conversion from a square image



**Fig. 11** ROC curve comparisons of the different algorithms

**Table 6 Comparison of proposed method with some of the state-of-the-art methods**

Comparison parameters	Ahmed et al. (2010)	Tang et al. (2008)	Monga and Mihcak (2007)	Khelifi and Jiang (2010)	Zhao et al. (2013)	Proposed method
Features used	Local	Global	Local	Local	Global and local features	Global and local features
Hash Length	7168 bits	320 bits	64 floating point numbers	250 bits	560 bits	848 bits
Robust against JPEG compression	Y	Y	Y	Y	Y	Y
Robust against rotation	NA	NA	NA	Y	Up to 5° only	All Arbitrary degree
Ability to detect small area forgery	Y	Y	Y	NA	Y	Y
Capability to find the counterfeit regions	Y	NA	NA	NA	Y	Y
Optimal TPR When FPR = 0	0.7843	0.8101	0.8213	0.8943	0.9157	0.9811
Optimal FPR When TPR = 1	0.3265	0.4972	0.3545	0.6427	0.2187	0.0012
Average time (s)	NA	0.9321	2.98	2.6	2.4	2.1

NA not applicable, Y Yes

into a secondary image. The other drawback is that the forged area cannot be accurately located under simultaneous rotation and malicious operation being carried out in the image.

The computational complexity of the proposed method and some of the state-of-the-art methods have been evaluated on a personal computer (with Intel core i7, 3.40 GHz, 8 GB RAM) running MATLAB 2013a. The comparison is shown in Table 6. It is observed that the proposed method is computationally efficient as compared to some of the existing methods. The reason for increased computational complexity as compared to the system proposed in Tang et al. (2008) may be due to the use of local features along with global features. Even though the computation complexity is not mentioned in Ahmed et al. (2010), it may be inferred that the major computation load is not because of wavelet transform but due to determining the decision threshold using various image processing operations.

**Conclusions**

In this paper, we have proposed a robust image hashing based on global features extracted through ring partition-PGNMF and local features that consist of position and texture information. The combination of global and local features is robust against the content-preserving operations, which has a desirable discriminative capability also. The local feature is useful to localize the forged area. The enhancement of the discriminating

capability of proposed hashing is due to PGNMF because it is good at learning part based representations. The PGNMF converges faster as compared to other NMFs. Experiments have been carried out on large image dataset to evaluate the performance of proposed hashing method. It is observed that proposed image hashing technique is robust to content preserving operations. It has also the capability to localize the forged area with sufficient accuracy.

The proposed algorithm may not be able to detect the color changes in the image. In future, the efficient color feature may be considered to detect the color forgery. Forged area localization under simultaneous rotation and malicious operations may also be explored. The other limitations such as reduction of hash length and computational complexity may be addressed during the future scope of work.

#### Abbreviations

DWT: discrete wavelet transform; DFT: discrete Fourier transform; MD: message digest; SHA: secure hash algorithm; DCT: discrete cosine transform; NMF: non-negative matrix factorization.

#### Authors' contributions

RKK and RHL carried out the study and implementation of Ring partitioning and PGNMF, participated in the sequence alignment and drafted the manuscript. BBR carried out the study and implementation of the spectral residual method. All authors read and approved the final manuscript.

#### Author details

<sup>1</sup> Department of Electronics and Communication Engineering, NIT Silchar, Silchar, Assam 788010, India. <sup>2</sup> Department of Electronics and Communication Engineering, NIT Mizoram, Aizawl, Mizoram 796012, India.

#### Acknowledgements

The authors would like to acknowledge people of Speech and Image Processing Laboratory, National Institute of Technology Silchar, India for providing support and necessary facilities for carrying out this work. The authors are thankful to Mr. Amarjit Roy, Ph.D. Scholar, Department of Electronics and Communication Engineering, for his constructive criticism and discussion.

#### Competing interests

The authors declare that they have no competing interests.

#### Funding

This study has not received any funding.

Received: 1 October 2015 Accepted: 3 November 2016

Published online: 21 November 2016

#### References

- Ahmed F, Siyal MY, Abbas VU (2010) A secure and robust hash-based scheme for image authentication. *Sig Process* 90(5):1456–1470
- Brunet JP, Tamayo P, Golub TR, Mesirov JP (2003) Metagenes and molecular pattern discovery using matrix factorization. *Proc Natl Acad Sci USA* 101(12):4164–4169
- Deselaers T, Keyzers D, and Ney H (2004) Features for image retrieval: a quantitative comparison. *Lecture Notes in Computer Science*, vol 3175. Springer, pp 228–236
- Donoho D, Stodden V (2004) When does non-negative matrix factorization give a correct decomposition into parts? In: Thrun S, Saul L, Scholkopf B (eds) *Advances in neural information processing systems* 16. MIT Press, Cambridge
- Fouad K, Jianmin J (2010) Analysis of the security of perceptual image hashing based on non-negative matrix factorization. *IEEE Signal Process Lett* 17(1):43–46
- Fridrich J, Goljan M (2000) Robust hash functions for digital watermarking. In: *Proceedings of the IEEE international conference information technology: coding*, pp 178–183
- Ground Truth Database (2008) <http://www.cs.washington.edu/research/imagetdatabase/groundtruth/>. Accessed 8 May 2008
- Hassan E, Chaudhury S, Gopal M (2012) Feature hashing, combination in kernel space for distance based image. *IEEE Trans Multimed* 14(4):1179–1195
- Hou X, Zhang L (2007) Saliency detection: a spectral residual approach. In: *Proceedings of the IEEE international conference on computer vision and pattern recognition*, Minneapolis, MN, pp 1–8
- Khelifi F, Jiang J (2010) Perceptual image hashing based on virtual watermark detection. *IEEE Trans Image Process* 19(4):981–994

- Kozat SS, Venkatesan R, Mihcak MK (2004) Robust perceptual image hashing via matrix invariants. In: Proceedings of the IEEE conference on image processing, pp 3443–3446
- Lee DD, Seung HS (2001) Algorithms for non-negative matrix factorization. In: Proceeding of the advances in neural information processing systems. MIT Press, Cambridge, Massachusetts, United States, pp 556–562
- Lefebvre F, Macq B, Legat JD (2002) RASH: radon soft hash algorithm. In: Proceedings of the European signal processing conference, pp 299–302
- Lei Y, Wang Y, Huang J (2011) Robust image hash in radon transform domain for authentication. *Signal Process Image Commun* 26(6):280–288
- Lin CJ (2007) Projected gradient method for nonnegative matrix factorization. *J Neural Comput* 19(10):2756–2779
- Lin CY, Chang SF (2001) A robust image authentication system distinguishing jpeg compression from malicious manipulation. *IEEE Trans Circuits Syst Video Technol* 11(2):153–168
- Lu W, Wu M (2010) Multimedia forensic hash based on visual words. In: Proceedings of the IEEE international conference on image processing, pp 989–992
- Lu CS, Hsu CY, Sun SW, Chang PC (2004) Robust mesh-based hashing for copy detection and tracing of images. In: Proceedings of the IEEE international conference on multimedia and expo, pp 731–734
- Lu W, Varna AL, Wu M (2010) Forensic hash for multimedia information. In: Proceedings of the SPIE, media forensics and security II, San Jose, CA
- Lv X, Wang ZJ (2009) Reduced-reference image quality assessment based on perceptual image hashing. In: Proceedings of the IEEE international conference on image processing, pp 4361–4364
- Lv X, Wang ZJ (2012) Perceptual image hashing based on shape contexts and local feature points. *IEEE Trans Inf Forensics Secur* 7(3):1081–1093
- Mishra M, Adhikary FLDMC (2013) Digital image tamper detection techniques: a comprehensive study. *Int J Comput Sci Bus Inform* 2:1–12
- Monga V, Mihcak MK (2007) Robust and secure image hashing via non-negative matrix factorizations. *IEEE Trans Inform Forensics Secur* 2(3):376–390
- Monga V, Banerjee A, Evans BL (2006) A clustering-based approach to perceptual image hashing. *IEEE Trans Inf Forensics Secur* 1(1):68–79
- Ouyang J, Coatrieux G, Shu H (2015) Robust hashing for image authentication using quaternion discrete Fourier transform and log-polar transform. *Digital Signal Proc* 41:98–109
- Paatero P, Tapper U (1994) Positive matrix factorization: a non-negative factor model with optimal utilization of error. *Environmetrics* 5:111–126
- Petitcolas FAP (2000) Watermarking schemes evaluation. *IEEE Signal Proc Mag* 17(5):1–4
- Qin C, Chang CC, Chen PY (2012) Self-embedding fragile watermarking with restoration capability based on adaptive bit allocation mechanism. *Sig Process* 92(4):1137–1150
- Qin C, Chen X, Ye D, Wang J, Sun X (2016) A novel image hashing scheme with perceptual robustness using block truncation coding. *Inf Sci* 361:84–99
- Roover CD, Vleeschouwer CD, Lefebvre F, Macq BK (2005) Robust video hashing based on radial projections of frames. *IEEE Trans Signal Process* 53(10):4020–4036
- Slaney M, Casey M (2008) Locality-sensitive hashing for finding nearest neighbors. *IEEE Signal Process Mag* 25(2):128–131
- Stamp M (2006) Information security: principles and practice, vol 2. Wiley, Hoboken
- Sun R, Zeng W (2014) Secure and robust image hashing via compressive sensing. *Multimedia Tools Appl* 70(3):1651–1665
- Swaminathan A, Mao Y, Wu M (2006) Robust and secure image hashing. *IEEE Trans Inform Forensics Secur* 1(2):215–230
- Tamura H, Mori S, Yamawaki T (1978) Textural features corresponding to visual perception. *IEEE Trans Syst Man Cybern* 8(6):460–472
- Tang Z, Wang S, Zhang X, Wei W, Su S (2008) Robust image hashing for tamper detection using non-negative matrix factorization. *J Ubiquitous Converg Technol* 2(1):18–26
- Tang Z, Wang S, Zhang X, Wei W, Zhao Y (2011) Lexicographical framework for image hashing with implementation based on DCT and NMF. *Multimed Tools Appl* 52(2–3):325–345
- Tang Z, Zhang X, Huang L, Dai Y (2013) Robust image hashing using ring-based entropies. *Signal Process* 93(7):2061–2069
- Tang Z, Zhang X, Zhang S (2014) Robust perceptual hashing based on Ring partition and NMF. *IEEE Trans Knowl Data Eng* 26(3):711–724
- Tang Z, Zhang X, Li X, Zhang S (2016a) Robust image hashing with ring partition and invariant vector distance. *IEEE Trans Inf Forensics Secur* 11(1):200–214
- Tang Z, Huang L, Zhang X, Lao H (2016b) Robust image hashing based on color vector angle and canny operator. *Int J Electronics Commun* 70:833–841
- USC-SIPI Image database (2007) <http://sipi.usc.edu/database/>. Accessed Feb 2007
- Venkatesan R, Koon SM, Jakubowski MH, Moulin P (2000) Robust image hashing. In: Proceedings of the IEEE international conference on image processing, pp 664–666
- Wang X, Pang K, Zhou X, Zhou Y, Li L, Xue J (2015) A visual model-based perceptual image hash for content authentication. *IEEE Trans Inf Forensics Secur* 10(7):1336–1349
- Wu MN, Lin CC, Chang CC (2007) Novel image copy detection with rotating tolerance. *J Syst Softw* 80(7):1057–1069
- Wu D, Zhou X, Niu X (2009) A novel image hash, algorithm resistant to print-scan. *Sig Process* 89(12):2415–2424
- Xiang S, Kim HJ, Huang J (2007) Histogram-based image hashing scheme robust against geometric deformations. In: Proceedings of the ACM multimedia and security workshop, New York, pp 121–128
- Xu W, Liu X, Gong Y (2003) Document clustering based on non-negative matrix factorization. In: Proceedings of the 26th annual international ACM SIGIR conference. ACM Press, pp 267–273. doi: 10.1145/860435.860485
- Yan CP, Pun CM, Yuan XC (2016) Multi-scale image hashing using adaptive local feature extraction for robust tampering detection. *Signal Proc* 121:1–16
- Zhao Y, Wang S, Zhang X, Yao H (2013) Robust hashing for image authentication using Zernike moments and local features. *IEEE Trans Inf Forensics Secur* 8(1):55–63