Published in IET Image Processing
Received on 21st December 2012
Revised on 16th April 2013
Accepted on 4th June 2013
doi: 10.1049/iet-ipr.2012.0717

ISSN 1751-9659

Forensic detection of image manipulation using the Zernike moments and pixel-pair histogram

Mahmood Shabanifard[1], *Mahrokh G. Shayesteh*[1,3], *Mohammad Ali Akhaee*[2]

[1]*Department of Electrical Engineering, Urmia University, Urmia, Iran*
[2]*School of Electrical and Computer Engineering, Faculty of Engineering, University of Tehran, Tehran, Iran*
[3]*Wireless Research Laboratory, ACRI, Electrical Engineering Department, Sharif University of Technology, Tehran, Iran*
E-mail: m.shayesteh@urmia.ac.ir

Abstract: Integrity verification or forgery detection of an image is a difficult procedure, since the forgeries use various transformations to create an altered image. Pixel mapping transforms, such as contrast enhancement, histogram equalisation, gamma correction and so on, are the most popular methods to improve the objective property of an altered image. In addition, fabricators add Gaussian noise to the altered image in order to remove the statistical traces produced because of pixel mapping transforms. A new method is introduced to detect and classify four various categories including original, contrast modified, histogram-equalised and noisy images. In the proposed method, the absolute value of the first 36 Zernike moments of the pixel-pair histogram and its binary form for each image in the polar coordinates are calculated, and then those features that yield the maximum between-class separation, are selected. Some other features obtained from Fourier transform are also utilised for more separation. Finally, support vector machine classifier is used to classify the input image into four categories. The experimental results show that the proposed method achieves high classification rate and considerably outperforms the previously presented methods.

1 Introduction

In recent decades, with growing digital imaging devices, digital images have been widely used in different aspects of society such as governmental, journalism, lawful or legal, and so on. One of the major problems of using digital images is that they can be easily modified using powerful editing image applications such as Photoshop; hence it is difficult to distinguish fake images from the genuine ones. The goals of forensic detection fall into the following categories [1]: (i) source classification, (ii) device identification, (iii) device linking groups objects according to their common source, (iv) integrity verification or forgery detection, (v) processing history recovery and (vi) anomaly investigation.

In this paper, we focus on the forgery detection and identify the process implemented on the input image such as contrast enhancement, histogram equalisation and noise addition. Most of the earlier works in this field have been devoted to disclose digital forgeries based on the incongruity in the chromatic aberration [2] by estimating the two-dimensional (2D) aberration-based model parameters and inconsistencies in the lighting angle of local regions in the images [3]. In [1], a new method was proposed based on photo response non-uniformity (PRNU) noise as the intrinsic fingerprint of imaging sensors along with maximum likelihood criterion to discriminate original images from the altered ones. In [4], the authors introduced a new method based on the absence or presence of the colour filter array interpolation-produced correlations of test images. These correlations would be destroyed when the original or genuine images were modified. Another method was presented in [5] to estimate the rotation angle and rescaling factor of local regions which are copy–moved in the altered image, based on the interpolation effect of the geometric operations. In [6], a new complex Zernike phase-based descriptor for region matching and local image representation has been introduced. However, this descriptor could not identify pixel value mapping-based transformations. In [7, 8] proposed a new scheme to detect the copy–move forgery in digital images using scale-invariant feature transform (SIFT) descriptors. The SIFT-based descriptors of an image are invariant to changes in scaling, rotation, illumination and so on. In [9], a new scheme was proposed based on the statistical traces produced by pixel value mapping applications called intrinsic fingerprints; however, the approach has two major drawbacks: (i) its performance degrades when the size of sub-images decreases; and (ii) the work is just organised into two-class method such as original and contrast-enhanced images or original and histogram-equalised images. The study in [10] exploited pattern noises which are the combination of PRNU and fixed pattern noise; camera identification algorithm was introduced based on extracting the reference pattern for each camera using correlation detector. For each camera under investigation, the authors first determine its reference pattern noise, which serves as a unique identification fingerprint. This is achieved by averaging the noise obtained from multiple images using a

denoising filter. To identify the camera from a given image, the reference pattern noise is considered as a spread spectrum watermark, whose presence in the image is established using a correlation detector. In [11], the author utilised PRNU parameter for forensic purposes. The work in [12] used double compression detection in JPEG images to detect manipulated images from a set of given images. In [13], a new method was introduced to detect image forgery based on the JPEG headers information such as quantisation table, Huffman codes, thumbnails and exchangeable image file (EXIF) format as the cameras signature. The authors in [14] calculated the second derivative of image in the frequency domain and used the normalised energy density criterion within different window sizes to detect resampled images from the original ones. They generated 19-D feature vectors employed to train a support vector machine (SVM) classifier. In [15], a new scheme was proposed to detect the copy–move forgery within an image. The performance of the previously proposed feature sets and the selected 15 most important feature sets were evaluated. The authors used a variety set of processing as the post-processing step such as matching, filtering, outlier detection, affine transformation estimation and so on. In [16] singular value decomposition (SVD) was exploited to decompose the RGB layer of an image into three rotation-invariant orthogonal matrices. In order to protect image against the forgery, 1D-cellular automata was used to generate the robust secret key. Then, the key was embedded into the spatial domain of another RGB layer to authenticate the original image.

In this paper, we present a new method to detect and distinguish pixel mapping transforms such as contrast enhancement, histogram equalisation and Gaussian noise addition to the images. In the first step, we divide an image into sub-images and then we remove low detail sub-images using entropy criterion to attain more accuracy in our classification. Next, the pixel-pair histogram of R, G and B components of each sub-image is calculated using different ordering patterns. The total pixel-pair histogram of each sub-image is calculated by peer-to-peer summation of the components of the obtained pixel-pair histograms. We also define binary pixel-pair histogram for each sub-image. Moreover, we calculate the absolute values of complex Zernike moments (ZMs) of the pixel-pair histograms of each sub-image in polar coordinates. Finally, SVM classifier is applied to classify the input feature set. In this study, we use Photoshop application to generate contrast-modified sub-images and MATLAB software to generate noisy and histogram-equalised ones. The results demonstrate the efficiency of the proposed method.

The rest of this paper is organised as follows: In Section 2, we explain the pixel-pair histogram of an image and complex ZM transform. In Section 3, we present the proposed method. The experimental results are provided in Section 4. Finally, Section 5 gives the concluding remarks.

2 Preliminaries

In this section, we briefly describe pixel-pair histogram and its corresponding binary form for an image. Then, we explain the F_{LAF} function of an image used in this paper and finally, the ZMs will be described.

2.1 Pixel-pair histogram and its binary form

In [9], the authors used the histogram of sub-images and proved that if the image is changed under pixel mapping transforms, the energy of high-frequency components of the histogram would increase. One disadvantage of using histogram as the first-order statistic is its low sensitivity to any pixel value manipulation. For example, suppose that one of the grey levels of an 8-bit image is removed due to using transform functions, then unlike the pixel-pair histogram that its effect is seen in the 255 components; only one component of histogram would be changed in the altered image.

On the other hand, pixel-pair histogram is frequently employed in the steganography [17] and steganalysis [18] applications. Therefore in the proposed method we use pixel-pair histogram instead of conventional histogram to obtain better performance in comparison with previous works in the same subject. In order to generate the pixel-pair histogram of an image, we convert the image into a 1D-vector using different ordering patterns such as row ordering, column ordering, zigzag ordering, shoe ordering and rotation one.

The pixel-pair histogram is an image of the size 256×256 (for grey-level images) where the intensity of each location (i, j) represents the number of times that the pixel pairs with the intensities i and j occurs in the generated vector. An example of the pixel-pair histogram of a typical image using column ordering pattern has been illustrated in Fig. 1.

Another type of pixel-pair histogram used in this article is referred to as the binary pixel-pair histogram. To calculate this histogram, the pixel-pair histogram values greater than one are clipped to one while the values equal to zero remain unchanged. Fig. 2 demonstrates the pixel-pair histogram and its binary form for an original image and its contrast-modified counterpart. We also employ binary pixel-pair histogram to identify the contrast enhancement fingerprint as seen in Fig. 2*e* in an altered image.

As observed from Figs. 2*c* and *f*, there exist many noisy-like points located out of the main diagonal and its neighbourhood in the binary form of the pixel-pair histogram. Hence, we use a mask to remove such noisy points in the binary form of pixel-pair histogram. The experimental results show that the performance of the SVM classifier improves when the mask is applied. The mask in fact is a white strip along the diagonal direction as shown in Fig. 2*e*. We have selected 100 sub-images randomly and calculated their pixel-pair histograms. Then, we used white strip with different widths to remove noisy points. Noting that the dimension of the pixel-pair histogram is 256×256, the results show that when the width of the white strip is selected equal to the square root of $((256/4)^2 + (256/4)^2)$, that is, 91 pixels, noisy points are well removed.

2.2 Logarithm of the absolute value of the image fast Fourier transform (FFT) (F_{LAF})

The 2D discrete Fourier transform (DFT) of an image function $f(x,y)$ of the size $M \times N$ pixels is defined as follows

$$F(u,v) = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x,y) \mathrm{e}^{-\mathrm{j}2\pi(ux/M+vy/N)} \tag{1}$$

In order to extract features from each pixel-pair histogram, we define F_{LAF} of the pixel-pair histogram as

$$F_{\mathrm{LAF}} = \log\left(1 + \mathrm{abs}\left(\mathrm{fft}\left(f(x,y)\right)\right)\right) \tag{2}$$

where fft(·) and abs(·) are the 2D-Fourier transform and the

doi: 10.1049/iet-ipr.2012.0717

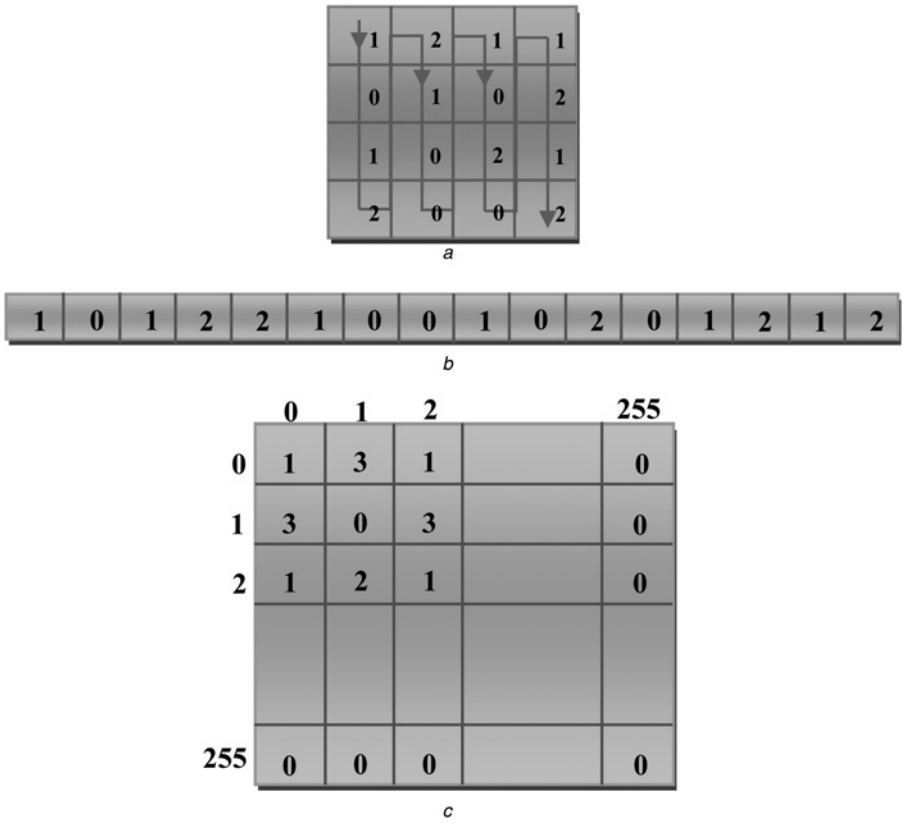

Fig. 1 *Example of pixel-pair histogram of a typical image using column ordering pattern*

a Typical image
b 1D-vector
c Pixel-pair histogram

absolute value operators, respectively. That is, first the magnitude of the FFT of the pixel-pair histogram is computed to obtain information about its frequency content. Then, we use logarithm operator to reduce the range of FFT values; for example, if the absolute FFT is in the range $[10^{-5}, 10^{5}]$, by taking the logarithmic function, the range falls within $[-5, 5]$. The value one is added to the magnitude to avoid log(0). The results show that the information in F_{LAF} can distinguish original images from the altered ones well.

2.3 Zernike moments

In most of the works, statistical moments, such as Geometric moments, Hermit moments, Gaussian–Hermit moments, ZMs and so on, are used as discriminating features. In recent decades, ZMs and their family have been widely employed in different applications such as object recognition [19], image retrieval [20], edge detection [21], image coding [22], image processing and analysing [23–25] and region descriptors [6, 26, 27], because they analyse the test images without considering changes in the position, size, viewing angle and orientation. The ZM coefficients are the outputs of the expansion of an image function into a complete orthogonal set of complex basis functions. Among many moment-based shape descriptors, ZM magnitude components are rotationally invariant and are more suitable for shape description.

To calculate the ZMs, the image (or region of interest) is first mapped to the unit disc using polar coordinates, where the centre of the image is the origin of the unit disc. Those pixels falling outside the unit disc are not used in the calculation. The mapping from Cartesian to the polar coordinates is

$$x = \rho\cos(\theta), \quad y = \rho\sin(\theta) \tag{3}$$

where ρ is the radius and θ is the angle

$$\rho = \sqrt{x^2 + y^2}, \quad \theta = \tan^{-1}\left(\frac{y}{x}\right) \tag{4}$$

The Zernike basis function $V_{nm}(\rho,\theta)$ is defined in polar coordinates over a unit circle as follows

$$V_{nm}(\rho,\theta) = R_{nm}(\rho)\mathrm{e}^{\mathrm{j}m\theta} \text{ for } \rho \le 1 \tag{5}$$

where n is the order and m is the repetition of $V_{nm}(\rho,\theta)$. Note that n and m are two non-negative integers where $|m| \le n$ and $n - |m|$ is even. $R_{nm}(\rho)$ is a radial polynomial which is given by

$$R_{nm}(\rho) = \sum_{s=0}^{(n-|m|)/2} (-1)^s \times \frac{(n-s)!}{s!\left(\left((n+|m|)/2\right)-s\right)!\left(\left((n-|m|)/2\right)-s\right)!}\rho^{n-2s} \tag{6}$$

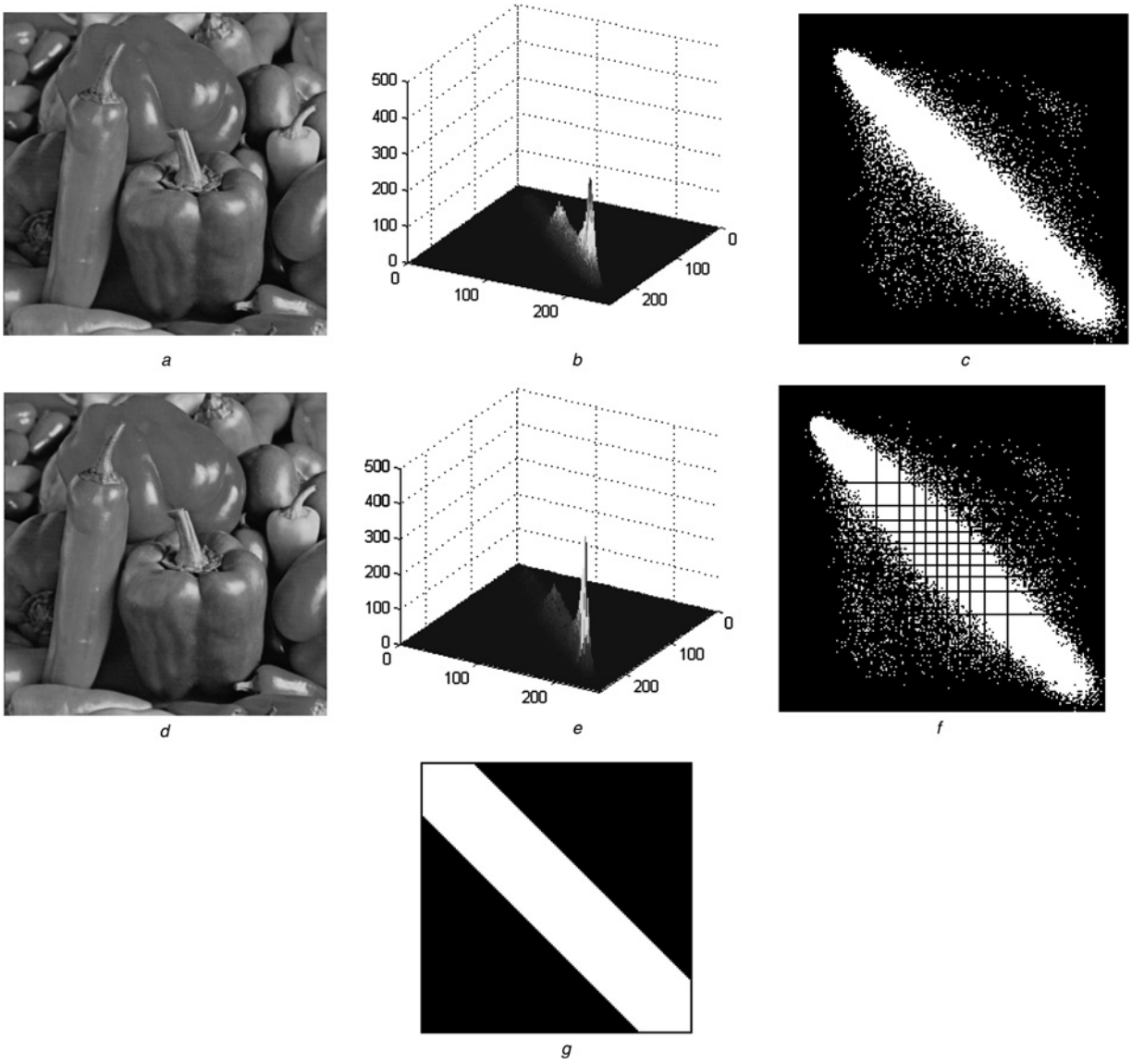

Fig. 2 *Pixel-pair histogram and its binary form*

a Original image
b Pixel-pair histogram
c Binary form of pixel-pair histogram
d Altered image by changing contrast to 20
e Pixel-pair histogram
f Binary form of the pixel-pair histogram
g Mask used to reject noisy like points

The set of basis functions is orthogonal, that is

$$\int_0^{2\pi}\int_0^1 V_{pq}(\rho,\theta)V_{nm}^*(\rho,\theta)\rho d\rho\ \mathrm{d}\theta = \frac{\pi}{n+1}\delta_{np}\delta_{mq} \qquad (7)$$

where

$$\delta_{ab} = \begin{cases} 1, & a = b \\ 0, & \text{else} \end{cases}$$

For a digital image function, the 2D-ZMs in the polar coordinates are given by

$$A_{nm} = \frac{n+1}{\pi}\sum_{(\rho,\ \theta)\in \text{ unit circle}}\sum f_p(\rho,\theta)V_{nm}^*(\rho,\theta) \qquad (8)$$

where $f_p(\rho,\theta)$ is the image function in polar coordinates.

The ZMs can be viewed as the responses of the image function to a set of quadrature-pair filters. In addition, repetition m indicates sector cycles of the function values along the azimuth angle, while n and m jointly specify a different number of annular patterns of the function.

3 Proposed method

Fig. 3 shows the schematic flowchart of the proposed algorithm. In the following, we explain different steps of the proposed method.

- *Step 1:* At first, the R, G and B components of each image are divided into sub-images of size 60×60 pixels. Then, those blocks that have higher entropy are considered for further processing; for example, the sub-images showing the sky or a smooth wall with a constant colour would be

doi: 10.1049/iet-ipr.2012.0717

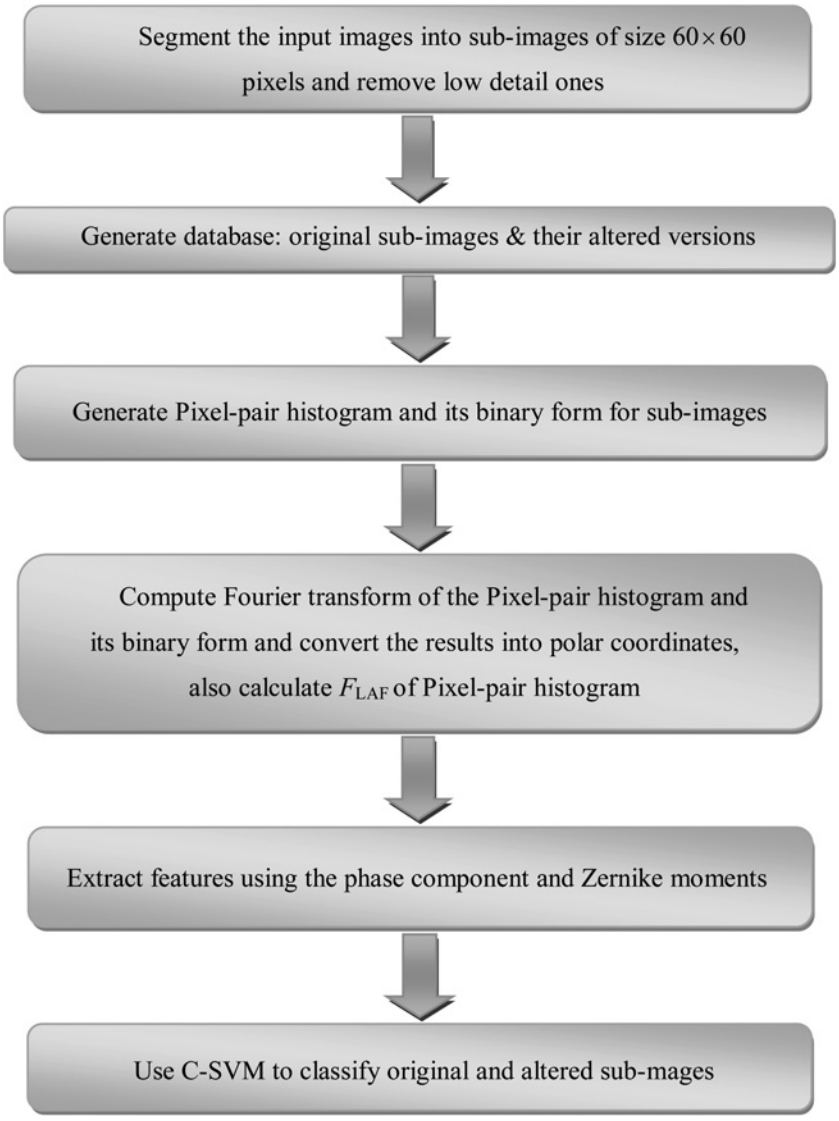

Fig. 3 *Block diagram of the proposed method*

rejected for performance improvement. We use the entropy criterion defined as

$$\text{Entropy} = -\sum_{i=1}^{M\times N} P(f_i)\log_2 P(f_i),\ \{f_i : i = 0,\ 1,\ \ldots,\ 255\} \tag{9}$$

where $M \times N$ is the size of the test sub-image and $P(f_i)$ is the probability of occurrence of the intensity f_i. In our experiment, blocks with entropy values equal to or greater than six are considered for feature extraction. It is worth mentioning that by examining the entropy values of 4.5 and 5, the SVM performance has shown slight degradation. This is why the threshold is set to six.

- *Step 2:* We generate our database including original, noisy, contrast-enhanced and histogram equalised sub-images. In order to compute the pixel-pair histograms of sub-images, *R*, *G* and *B* components of each sub-image are considered separately.
- *Step 3:* Different ordering patterns, as explained before, are exploited to calculate the pixel-pair histograms of each sub-image. We estimate the total pixel-pair histogram by peer to peer summation of the components of the calculated pixel-pair histograms. We then compute the binary form of the resulted pixel-pair histogram.
- *Step 4:* The F_{LAF} function (the logarithm of the absolute value of the Fourier transform pixel-pair histogram and its binary form of each sub-image) using (2) is computed. Then, we convert the results from Cartesian coordinates into the polar coordinates.
- *Step 5*: As our main features, specific ZMs of F_{LAF} function are calculated by using (8). The other feature is obtained from the distribution of absolute value of the Fourier transform of the pixel-pair histogram and its binary form.
- *Step 6:* In the final step, we apply C-SVM (standard soft-margin SVM) classifier with RBF kernel to classify the original images from their altered versions.

In the following, more explanations about feature selection (step 5) will be presented.

3.1 Dataset generation and feature selection

We use UCID image database [28] as unaltered or original images database. We extract R, G and B components of

each original image as the grey-scale images. Next, the three components R, G and B are divided into sub-images of size 60×60 pixels and smooth sub-images are rejected according to our entropy criterion. We select 1000 high-entropy sub-images randomly as the original signals. We generate contrast-enhanced sub-images by using Adobe Photoshop CS5 application and change the contrast values of original sub-images to +40. Further, we use MATLAB software to add zero mean white Gaussian noise with variances 0.01 to the original sub-images. Finally, we generate the histogram-equalised sub-images via MATLAB software using the following equation

$$H(i) = \text{round}\left[255 \sum_{n=0}^{i} \frac{h(n)}{M \times N}\right] \tag{10}$$

where $H(i)$ is the histogram-equalised sub-image with the intensity i, $h(n)$ is the histogram value of the input sub-image with the intensity n, and $M \times N$ is the total number of pixels in the original sub-image. Consequently, our dataset has totally 4000 sub-images (1000 from each of four categories). This set is used to find the proper features of each category which result in enough separation and high performance. As stated in step 5, one of the features employed in the proposed method is the distribution of the absolute value of Fourier transform of the pixel-pair histogram and its binary form in the polar coordinates as a function of angle in the range 0° to 180° and maximum radius which is based on the size of the pixel-pair histogram. For example, for the image of size $M \times N$, the radius will be $\min\{M/2, N/2\}$. Here, the maximum radius is set to 128. If we show the absolute value of the Fourier transform of an image by $S(\rho, \theta)$, then we calculate the distribution S_{ang} as follows

$$S_{\text{ang}} = S(\theta) = \sum_{r=1}^{R_o} S_r(\theta), \quad 0 \le \theta < \pi \tag{11}$$

where $S_r(\theta)$ is the 1D-function of $S(\rho, \theta)$ for each frequency r and R_o is the radius of the circle centred at the origin depending upon the size of pixel-pair histogram (more details are explained in [29]). In the following, we will discuss the relevant features of each class.

3.2 Specific features of contrast-enhanced sub-images

In order to select features to detect contrast-enhanced images, we calculate 36 ZMs of the F_{LAF} function (2) of the pixel-pair histogram and its binary form from the generated dataset. Then, we generate feature vectors where the first 36 elements of each feature vector are calculated from pixel-pair histogram and the next 36 elements are computed from its binary form. Therefore, our primary feature vectors have 72 dimensions. Since the values of many dimensions of the feature vectors are about 10^5 while the range of other dimensions is far less than it, we normalise the elements of each feature vector. Our experimental results show that the PDF of each dimension has the Gaussian distribution; therefore the following equation for normalisation can be used

$$D_{\text{norm}}(i) = \frac{D(i) - \mu_{D(i)}}{\delta_{D(i)}} \tag{12}$$

where $D(i)$ is the ith dimension of the feature vector $\boldsymbol{D}$, $\mu_{D(i)}$ and $\delta_{D(i)}$ are the mean and standard deviation of $D(i)$, respectively and $D_{\text{norm}}(i)$ is the normalised value of $D(i)$.

Finally, we apply mutual information technique to reduce the dimension of the feature vectors. For this purpose, we apply minimum redundancy-maximum relevance (mRMR) method [30] to the feature vectors. Fig. 4 shows the SVM performance against the selected dimensions by mRMR method. It is observed that the SVM performance is more than 98.5%. Considering the figure, we select the dimension 10 which achieves the accuracy 99.5%. Table 1 shows the 10 best moments indices and the equivalent ZMs of the contrast-enhanced class obtained by mRMR method. We have used the notations 'dir' and 'bin' in tables which stand for using pixel-pair histogram directly and using its binary form, respectively.

The other features are obtained from the curve of contrast-enhanced sub-images (11), which is calculated by the absolute value of the Fourier transform of the binary form of pixel-pair histogram. To obtain the other features of this category, we have computed S_{ang}s of 50 randomly selected images. Figs. 5*a* and *b* show the S_{ang} curves of 15 original sub-images and their contrast-modified counterparts for pixel-pair histogram and its binary form, respectively. As seen in Fig. 5*a*, there is no difference between the S_{ang} curves of original and altered version in the case of the pixel-pair histogram. However, from Fig. 5*b*, we observe that there are three noticeable peaks around $\theta = 0°$, 90° and 180° in the binary pixel-pair histogram of the altered

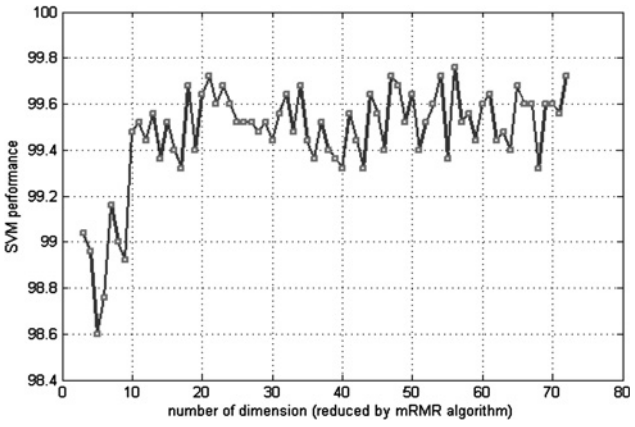

Fig. 4 *SVM performance (%) against the selected dimensions by mRMR method for contrast-enhanced sub-images*

Table 1 Ten best moment indices and equivalent ZMs for contrast-enhanced images

Moment index	58	24	70	59	54	60	57	62	22	38
ZMs	$n=6$ $m=4$ bin	$n=10$ $m=4$ dir	$n=10$ $m=8$ bin	$n=8$ $m=4$ bin	$n=5$ $m=3$ bin	$n=10$ $m=4$ bin	$n=4$ $m=4$ bin	$n=7$ $m=5$ bin	$n=6$ $m=4$ dir	$n=2$ $m=0$ bin

doi: 10.1049/iet-ipr.2012.0717

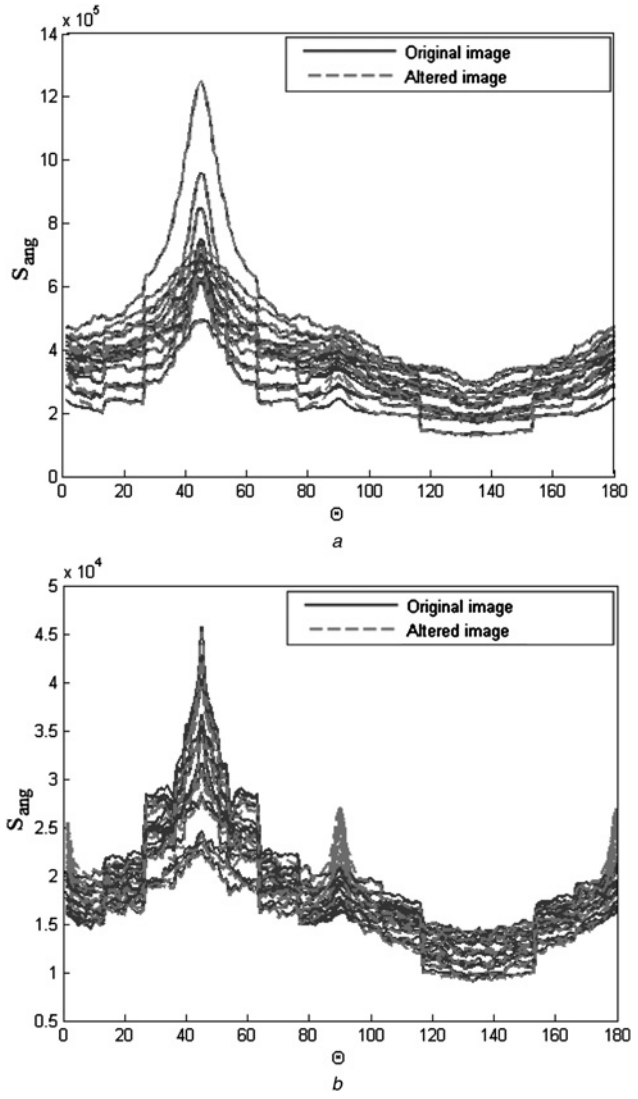

Fig. 5 *Plot of S_{ang} curves*

a Pixel-pair histogram
b Binary form of pixel-pair histogram for 15 original sub-images and their modified contrast versions

sub-images. Therefore, considering Fig. 5*b*, we define two new features. They are obtained from calculating the sum of peak values of S_{ang} at $\theta = 0°$, 90° and 180° and the sum of variances of S_{ang} in the intervals [0:10], [80:100] and [170:180] degrees. Note that we cannot use the peak around $\theta = 45°$ to extract a new feature because at $\theta = 45°$, the S_{ang} curve has similar behaviour for the original sub-images and their contrast-modified counterparts.

As a consequence, we use 12 features for this category (ten features belong to ZMs of the pixel-pair histogram and its binary form and two features obtained from S_{ang} of binary form of the pixel-pair histogram). These features are able to classify original sub-images from the contrast-enhanced ones.

3.3 Specific features of histogram-equalised sub-images

The features extracted in part B do not classify histogram-equalised images from other classes. Thus, using the same procedure of part B, we select a new set of ZMs according to Table 2 based on the maximum class separation they provide. Our results showed that the classification performance is always 100% even with one feature. Thus, we select one dimension to distinguish histogram-equalised dataset from other classes. We did not use S_{ang} curve to define a new feature set for this category since the ZM is sufficient to distinguish this class from the genuine one.

3.4 Specific features of noisy sub-images

Fabricators may add noise to remove the trace of pixel mapping transforms in their works. Moreover, the introduced features in parts A and B are not able to isolate noisy sub-images from other categories. This leads us to find a new feature set. For this purpose, we define a set of ZMs, which result in enough separation between the noisy dataset and original one. In this experiment, we added Gaussian noises with zero mean and variance 0.01 to the original images. The number of dimensions selected by mRMR method is four. Table 3 shows the four best ZMs to identify this category from the original class.

Another feature is obtained by computing the Fourier transform of the pixel-pair histogram in our dataset and then calculating the absolute value of the summation of the normalised real parts of its Fourier coefficients. Fig. 6 shows this feature for 100 original and noisy sub-images with variance of 0.02. Thus, we have totally five features for noisy sub-images dataset. Note that we did not use S_{ang} curve to define a new set feature for noisy category because the original and noisy S_{ang} curves have shown similar behaviours.

3.5 Final feature vector

To generate the final feature vector, we extract the mentioned features mentioned in parts B, C and D for each of pixel-pair histogram and its binary form of all sub-images. Therefore, the total feature vector has 18 dimensions, where 12 features correspond to those calculated for contrast-enhanced dataset, one feature is computed as in the histogram-equalised dataset, and the last five dimensions are those used for noisy dataset.

As stated, we have defined two feature sets which are based on the ZMs and features obtained from Fourier transform. It will be shown in the next section that the performance increases using these two kinds of features.

4 Experimental results

In this section, we evaluate the performance of the proposed method. The dataset is generated in the same way mentioned

Table 2 Ten best moment index and equivalent ZMs for histogram-equalised images

Moment index	55	17	41	68	62	40	53	6	61	59
ZMs	$n=7$ $m=3$ bin	$n=3$ $m=3$ dir	$n=8$ $m=0$ bin	$n=9$ $m=7$ bin	$n=7$ $m=5$ bin	$n=6$ $m=0$ bin	$n=3$ $m=3$ bin	$n=10$ $m=0$ dir	$n=5$ $m=5$ bin	$n=8$ $m=4$ bin

Table 3 Four best moment indices and equivalent ZMs for noisy images

Moment index	5	58	16	6
ZMs	$n=8$ $m=0$ dir	$n=6$ $m=4$ bin	$n=10$ $m=2$ dir	$n=10$ $m=0$ dir

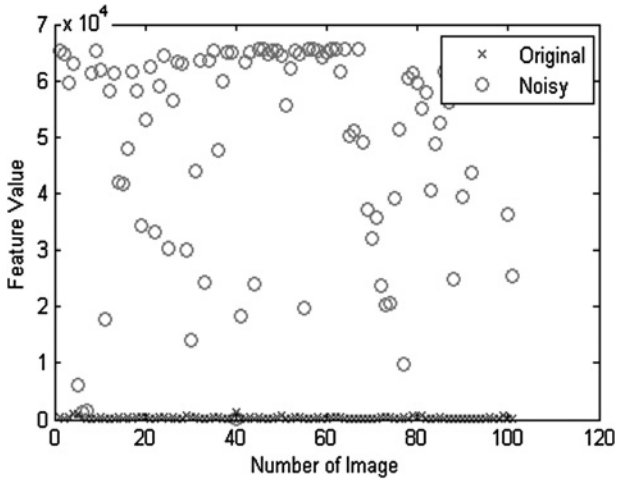

Fig. 6 *Values of feature obtained from the Fourier transform of 100 original sub-images and their noisy versions for pixel-pair histogram*

in the previous section. However, we select 2500 of them randomly as the original dataset and thanks to Adobe Photoshop CS5 application; the contrast values of original sub-images are converted into −50, −40, −20, −10, −5, +5, +10, +20, +40, +60 and +100. Moreover, we use MATLAB software to add zero mean white Gaussian noise with variances 0.001, 0.002, 0.005, 0.01 and 0.02 to the original sub-images and contrast-enhanced ones. We also generate histogram-equalised sub-images considering (10) in MATLAB environment. Finally, we generate feature vectors of each class as mentioned in Section 3.

4.1 Classification results

In order to evaluate the performance of the proposed method, we apply SVM classifier. Although SVMs are originally designed as binary classifiers, there exist various extensions to enable SVMs to handle more than two classes. Multiclass SVM classifiers can be roughly divided into two groups, all-together methods and the methods based on the binary classifiers. In this work, for multiclass classification, we use One-against-One (OvO) method which is widely used in the literature. This method employs $\binom{N_c}{2}$ binary SVMs for each pair of classes, where N_c is the number of classes. During the classification, the feature vectors are presented to all $\binom{N_c}{2}$ binary classifiers and the histogram of the outputs is evaluated. The class corresponding to the maximum value of the histogram is selected as the target class. If there are two or more classes with the same number of votes, one of the classes is randomly chosen. In multiclass SVM, we also use C-SVM with radial basis function (RBF) kernel as binary classifier bases, which is a commonly used kernel.

Here, we employ LIBSVM library with RBF kernel as $\exp\left(-\lambda||D^{(l)}-D^{(k)}||^2\right)$ for kernel function where $\lambda>0$ is the Gaussian kernel width. The developer of LIBSVM library suggested users to try the RBF kernel first and if RBF is used with model selection, then there is no need to consider the linear kernel. The kernel matrix using sigmoid may not be positive definite and in general its accuracy is not better than RBF.

In the first experiment, the feature vectors of all datasets are divided into two groups; the first group is exploited to train the SVM classifier that contains 1500 feature vectors for each dataset and the second group is used to test the trained SVM that includes 1000 feature vectors for each category. We train the SVM classifier with feature vectors from the original dataset, noisy dataset with the variance of 0.01, histogram-equalised dataset and contrast-enhanced categories with the contrast values of 10, 100 and −10 (totally 4500 vectors for contrast-enhanced class). In order to test the classifier performance in contrast-enhanced case, we use a variety of contrast-enhanced images with parameters −50, −40, −20, −10, −5, 5, 10, 40, 60 and 100.

Table 4 demonstrates the average classification accuracy when 1000 different test images from each dataset are used. The performance is evaluated in the cases that the parameter of the contrast-enhanced category varies in the test phase. From the table, it is observed that the accuracy of the proposed method is more than 93% for all contrast values expect the contrast value of −5 which is about 80%. The reason for this rather less performance is that the contrast enhancement pattern is drawn to the both sides of the main diagonal of the binary pixel-pair histogram for the contrast value equal to −5; therefore it is difficult to recognise the generated contrast pattern. From Table 4, it is also noticed that the classifier attains high accuracy when the test sub-images of the contrast-enhanced class are altered by different contrast parameters than those used in the training stage. That is, the SVM is trained with three contrast values, while it is tested with the same and different contrast values.

We have also used the confusion matrix to show the performance of the proposed method. The confusion matrix is a specific table layout that allows visualisation of the performance of an algorithm. Each column of the matrix represents the instances in a predicted class, whereas each row represents the instances in an actual class. Table 5 demonstrates the confusion matrix of four categories when in the test phase different values of contrast-enhanced images are applied. Moreover, the classification rates of different classes are presented. It is observed that the proposed method can recognise different categories with high accuracy. It should be noted that although we have trained the SVM classifier with three parameters of contrast-enhanced images, the classifier can recognise the contrast-enhanced images with parameters different from those used in the training phase.

To justify the reason of employing two sets of features simultaneously, we show the effect of each set on the performance separately. To this end, we train the SVM classifier using just ZMs with the same manner mentioned before, then we test the SVM classifier with for original dataset, noisy dataset with the variance of 0.01, histogram-equalised dataset and contrast-enhanced images with different contrast values. The accuracy rates for the contrast values of −10, 5, 10 and 20 were obtained as 96.8, 90,

doi: 10.1049/iet-ipr.2012.0717

Table 4 Average performance (%) of the proposed method for four categories against different contrast values used in test phase

Value of contrast-enhanced data	Classification accuracy, %
−50	99.00
−40	99.10
−20	98.45
−10	99.15
−5	80.01
+5	93.00
+10	97.65
+20	99.03
+40	99.10
+60	99.01
+100	99.15
average	96.60

97.50 and 97.45%, respectively. We also trained the classifier with the other feature set and tested as stated above. The performance of the classifier for the above contrast values was about 97.75, 94.55, 88.80 and 97.72%. Although the accuracies in the case of using two set sets of features noting Table 4 were attained as 99.15, 93, 97.65 and 99.03%, respectively, the results indicate that the performance improves when both sets of features are exploited.

As another experiment, we examine the effect of the variance of Gaussian noise on the performance of the proposed method. For this purpose, we generate noisy sub-images from the original ones. Here, we train the SVM classifier with images from the original dataset, noisy dataset with the variances of 0.01 and 0.005, histogram-equalised dataset and contrast-enhanced categories with the contrast values of 10, 100 and −10. We then test our dataset consisting of sub-images with noise variances of 0.001, 0.002, 0.005, 0.01 and 0.02 and different contrast values. Table 6 shows the average performance of the classifier. It is observed that the proposed method achieves high accuracy. Moreover, the performance of contrast-enhanced and noisy image categories are listed separately in this table.

We have also evaluated the performance of the trained SVM, in the case that Gaussian noise is added to the contrast-modified sub-images instead of adding to the original sub-images. That is, we generate the new noisy dataset by adding Gaussian noise with different variances to the contrast-enhanced sub-images and then tested the trained SVM by the newly generated noisy dataset. Note that we used previous trained SVM in this experiment. As expected, the SVM classifier recognises and classifies noisy contrast-enhanced sub-images as noisy sub-images; that is, the effect of additive noise on any altered type of image is like its effect on the original image. Table 7 demonstrates the performance of the classifier in identifying each category. It is observed that the proposed method achieves high classification rates.

4.2 Performance comparison

Here, we compare the new method with the scheme recently proposed in [9] in which the contrast-enhanced dataset is generated as follows

$$f_c(i) = \text{round}\left[255(i/255)^{\gamma}\right] \quad (13)$$

Table 5 Confusion matrix of datasets and classification accuracy for different contrast values used in the test phase

Target class	Output class				Accuracy, %
	Original	Noisy	Cont. en	Hist. eq	
(a) −50					
original	985	6	9	0	98.50
noisy	10	990	0	0	99.00
cont. en	11	0	989	0	98.90
hist. eq	0	0	0	1000	100.00
(b) −40					
original	985	6	9	0	98.50
noisy	10	990	0	0	99.00
cont. en	3	2	995	0	99.50
hist. eq	0	0	0	1000	100.00
(c) −20					
original	985	6	9	0	98.50
noisy	10	990	0	0	99.00
cont. en	44	2	954	0	95.40
hist. eq	0	0	0	1000	100.00
(d) −10					
original	985	6	9	0	98.50
noisy	10	990	0	0	99.00
cont. en	2	0	998	0	99.80
hist. eq	0	0	0	1000	100.00
(e) +10					
original	985	6	9	0	98.50
noisy	10	990	0	0	99.00
cont. en	26	0	974	0	97.40
hist. eq	0	0	0	1000	100.00
(f) +20					
original	985	6	9	0	98.50
noisy	10	990	0	0	99.00
cont. en	0	0	1000	0	100.00
hist. eq	0	0	0	1000	100.00
(g) +40					
original	985	6	9	0	98.50
noisy	10	990	0	0	99.00
cont. en	1	0	999	0	99.90
hist. eq	0	0	0	1000	100.00
(h) +60					
original	985	6	9	0	98.50
noisy	10	990	0	0	99.00
cont. en	0	0	1000	0	100.00
hist. eq	0	0	0	1000	100.00

where f_c is the contrast-enhanced image and gamma (γ) is a positive constant. The authors in [9] used two classes, including original and contrast-enhanced sub-images.

Here, for fair comparison, we generated sub-images of size 50×50 and 20×20 pixels and modified them using (13) by different gamma values as used in [9]. It is noteworthy that the sub-images were selected from the total generated sub-images which have entropy values equal to or greater than five. We test the two methods under our dataset.

As stated before, the performance of the proposed method is almost the same for the entropy value equal to 4.5. The total number of sub-images is equal to 26 352 for the block size of 50×50 pixels and by choosing the entropy value equal to six, the total remained sub-images will be 24 285, that is, 2067 sub-images are removed from dataset.

Now, we examine the proposed method and the method of [9] using the new generated datasets. We train SVM classifier with the two mentioned categories to have two-class problem. We use receiver operating characteristic (ROC) curves for

Table 6 Performance (%) of the proposed method for different values of contrast and noise variances

			Variance of Gaussian noise									
			0.001		0.002		0.005		0.01		0.02	
			Contrast	Noise	Contrast	Noise	Contrast	Noise	Contrast	Noise	Contrast	Noise
contrast enhancement value	100	category performance	100	14.80	100	49.2	100	97.7	100	97.7	100	99.7
		average performance	78.43		87.03		99.15		99.15		99.65	
	60	category performance	100	14.80	100	49.2	100	97.7	100	97.7	100	99.7
		average performance	78.43		87.03		99.15		99.15		99.65	
	40	category performance	100	14.80	100	49.2	100	97.7	100	97.7	100	99.7
		average performance	78.43		87.03		99.15		99.15		99.65	
	20	category performance	99.60	14.80	99.60	49.2	99.60	97.7	99.60	97.7	99.60	99.7
		average performance	78.33		86.93		99.05		99.05		99.55	
	10	category performance	93.40	14.80	93.40	49.2	93.40	97.7	93.40	97.7	93.40	99.7
		average performance	76.78		85.38		97.50		97.50		98.00	
	5	category performance	74.20	14.80	74.20	49.2	74.20	97.7	74.20	97.7	74.20	99.7
		average performance	72.01		80.58		92.7		92.7		93.2	
	−5	category performance	20.60	14.80	20.60	49.2	20.60	97.7	20.60	97.7	20.60	99.7
		average performance	58.58		67.18		79.3		79.3		79.80	
	−10	category performance	100	14.80	100	49.2	100	97.7	100	97.7	100	99.7
		average performance	78.43		87.03		99.15		99.15		99.00	
	−20	category performance	97.00	14.80	97.00	49.2	97.00	97.7	97.00	97.7	97.00	99.7
		average performance	77.68		86.28		98.40		98.40		99.60	
	−40	category performance	99.80	14.80	99.80	49.2	99.80	97.7	99.80	97.7	99.80	99.7
		average performance	78.38		86.98		99.10		99.10		99.48	
	−50	category performance	99.30	14.80	99.30	49.2	99.30	97.7	99.30	97.7	99.30	99.7
		average performance	78.30		86.85		98.98		98.98		99.20	

doi: 10.1049/iet-ipr.2012.0717

Table 7 Performance (%) of the proposed method for different values of contrast and noise variance when noise is added to the contrast-enhanced sub-images

			Variance of Gaussian noise							
			0.002		0.005		0.01		0.05	
			Contrast	Noise	Contrast	Noise	Contrast	Noise	Contrast	Noise
contrast enhancement value	100	category performance	100	59.50	100	89.70	100	98.10	98.50	98.50
		average performance	89.60		97.15		99.25		99.35	
	60	category performance	100	59.50	100	89.70	100	98.10	98.50	98.50
		average performance	89.60		97.15		99.25		99.35	
	40	category performance	100	59.50	100	89.70	100	98.10	98.50	98.50
		average performance	89.60		97.15		99.25		99.35	
	20	category performance	99.60	59.50	99.60	89.70	99.60	98.10	98.50	98.50
		average performance	89.50		97.05		99.15		99.25	
	10	category performance	93.40	59.50	93.40	89.70	93.40	98.10	98.50	98.50
		average performance	87.95		95.50		97.60		97.70	
	5	category performance	74.20	59.50	74.20	89.70	74.20	98.10	98.50	98.50
		average performance	83.15		90.70		92.80		93.00	
	−5	category performance	20.60	59.50	20.60	89.70	20.60	98.10	98.50	98.50
		average performance	69.80		77.30		79.40		79.50	
	−10	category performance	100	59.50	100	89.70	100	98.10	98.50	98.50
		average performance	89.60		97.15		99.25		99.35	
	−20	category performance	97.00	59.50	97.00	89.70	97.00	98. 10	98.50	98.50
		average performance	88.85		96.4		98.50		98.60	
	−40	category performance	99.80	59.50	99.80	89.70	99.80	98.10	98.50	98.50
		average performance	89.55		97.10		99.20		99.30	
	−50	category performance	99.3	59.50	99.3	89.70	99.3	98.10	98.50	98.50
		average performance	89.43		96.98		99.08		99.18	

performance comparison. A ROC space is defined by false positive rate (FPR) and true positive rate (TPR) as x and y axes, respectively, which depicts relative trade-offs between true positive (benefits) and false positive (costs). Since TPR is equivalent with sensitivity and FPR is equal to 1-specificity, the ROC graph is sometimes called the sensitivity against 1-specificity plot. Each prediction result or instance of a confusion matrix represents one point in the ROC space. The best possible prediction method would yield a point in the upper left corner or coordinate (0, 1) of the ROC space, representing 100% sensitivity (no false negatives) and 100% specificity (no false positives). The (0, 1) point is also called the perfect classification.

We have evaluated the method of [9] on our dataset for sub-images of size 50×50 and 20×20 pixels. Fig. 7*a* depicts the ROC curve for gamma (γ) equal to 0.9. Moreover, Fig. 7*b* demonstrates the ROC plot of the proposed method on the same database for block sizes of 50×50 and 20×20 pixels and the same gamma (γ). It is observed that our method classifies the two datasets with high accuracy. For instance, as the performance comparison, considering Fig. 7*a*, for the probability of detection P_d (true positive rate) about 55% using the method of [9], the probability of false alarm P_{fa} (FPR) is about 5% for blocks of size 50×50 pixels. These values will be $P_d \simeq 50\%$ and $P_{fa} \simeq 22\%$ when the size of sub-images reduces to 20×20 pixels. On the other hand, in the proposed method, as shown in Fig. 7*b* for the same gamma value, we achieve $P_d > 99.5\%$ and $P_{fa} < 5\%$ for block sizes of 50×50 pixels and $P_d > 75\%$ and $P_{fa} < 5\%$ for block sizes of 20×20 pixels. That is, the new method significantly outperforms the method of [9]. Another point

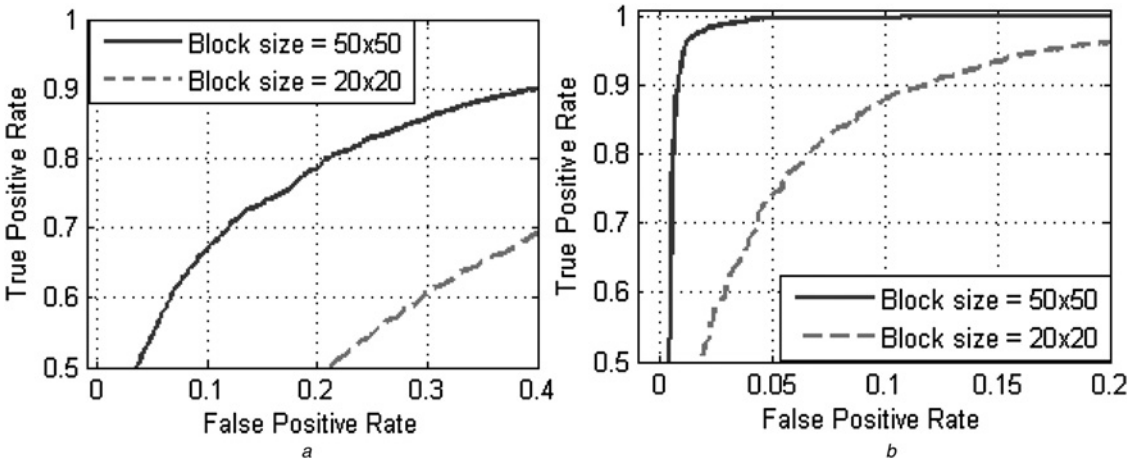

Fig. 7 *ROC curves (true positive rate against FPR) obtained using the method of [9] evaluated on our dataset and the proposed method for $\gamma = 0.9$*

a Method of [9]
b Proposed method

is that the proposed scheme can classify four classes whereas the method of [9] is a two-class problem.

5 Conclusion

Image forgery using pixel mapping transform is a very common forgery manner in order to make the same lighting conditions in the altered images. Researchers have proposed several methods to detect and specially locate the modified regions on altered images. However, many of them were not successful to detect and identify the pixel mapping transform utilisation in the test images. We have proposed a new method to detect and identify the type of used pixel mapping transform on the altered images with high accuracy. To evaluate the performance of the proposed method, we generated four different datasets namely original, histogram-equalised, contrast-enhanced and noisy datasets. For each category, specific features are extracted from the ZMs of the F_{LAF} function of pixel-pair histogram and its binary form. Besides, some features are generated from the distribution of the absolute value of the Fourier transform of the pixel-pair histogram and its binary form in the polar coordinates as a function of angle in the range of 0–180°. Simulation results indicate that the introduced method can classify the four categories with high accuracy and outperforms the state-of-the-art techniques. In the future work, we focus on finding new features to detect and identify the geometric transform such as rotation, rescaling and so on, and then the corresponding robust detection method will be investigated.

6 References

1 Chen, M., Fridrich, J., Goljan, M., Lukas, J.: 'Determining image origin and integrity using sensor noise', *IEEE Trans. Inf. Forensics Sec.*, 2008, **3**, (1), pp. 74–90

2 Johnson, M.K., Farid, H.: 'Exposing digital forgeries through chromatic aberration'. Proc. ACM Multimedia and Security Workshop, Geneva, Switzerland, 2006, pp. 48–55

3 Johnson, M.K., Farid, H.: 'Exposing digital forgeries in complex lighting environments', *IEEE Trans. Inf. Forensics Sec.*, 2007, **2**, (3), pp. 450–461

4 Popescu, A.C., Farid, H.: 'Exposing digital forgeries in color filter array interpolated images', *IEEE Trans. Signal Process.*, 2005, **53**, (10), pp. 3948–3959

5 Wei, W., Wang, S., Zhang, X., Tang, Z.: 'Estimation of image rotation angle using interpolation-related spectral signatures with application to blind detection of image forgery', *IEEE Trans. Inf. Forensics Sec.*, 2010, **5**, (3), pp. 507–517

6 Chen, Z., Sun, S.K.: 'A Zernike moment phase-based descriptor for local image representation and matching', *IEEE Trans. Image Process.*, 2010, **19**, (1), pp. 205–219

7 Huang, H., Guo, W., Zhang, Y.: 'Detection of copy-move forgery in digital images using SIFT algorithm'. Proc. IEEE Pacific-Asia Workshop on Computational Intelligence and Industrial Application, 2008

8 Pan, X., Lyu, S.: 'Detecting image region duplication using SIFT features'. Proc. ICASSP, 2010, pp. 1706–1709

9 Stamm, M.C., Ray Liu, K.J.: 'Forensic detection of image manipulation using statistical intrinsic fingerprints', *IEEE Trans. Inf. Forensics Sec.*, 2010, **5**, (3), pp. 492–506

10 Lukás, J., Fridrich, J.J., Goljan, M.: 'Digital camera identification from sensor pattern noise', *IEEE Trans. Inf. Forensics Sec.*, 2006, **1**, (2), pp. 205–214

11 Fridrich, J.: 'Digital image forensic using sensor noise', *IEEE Trans. Signal Process.*, 2009, **26**, (2), pp. 26–37

12 Mahdian, B., Saic, S.: 'Detecting double compressed JPEG images'. Proc. Third Int. Conf. Imaging for Crime Detection and Prevention (ICDP-09), 2009

13 Kee, E., Johnson, M.K., Farid, H.: 'Digital image authentication from JPEG headers', *IEEE Trans. Inf. Forensics Sec.*, 2011, **6**, (3), pp. 1066–1075

14 Cox, I.J., Doerr, G.: 'Normalized energy density-based forensic detection of resampled images', *IEEE Trans. Multim.*, 2012, **14**, (3), pp. 536–545

15 Riess, C., Jordan, J., Riess, C., Angelopoulou, E.: 'An evaluation of popular copy-move forgery detection approaches', *IEEE Trans. Inf. Forensics Sec.*, 2012, **7**, (6), pp. 1841–1854

16 Pahlavan Tafti, A., Rohani, F., Moghaddasifar, M.A.: 'RGB digital image forgery detection using singular value decomposition and one dimensional cellular automata'. 15th IEEE Int. Conf. Computing Technology and Information Management (ICCM), 2012, pp. 483–488

17 Cancelli, G., Doerr, G., Cox, I.J., Barni, M.: 'Detection of ±1 LSB steganography based on the amplitude of histogram local extrema'. 15th IEEE Int. Conf. Image Processing (ICIP), 2008, pp. 1288–1291

18 Lan, D.Q.: 'The blind detection of information hiding in color image'. Proc. Second Int. Conf. Computer Engineering and Technology (ICCET), 2010, vol. 7, pp. 346–348

19 Shutler, J.D., Nixon, M.S.: 'Zernike velocity moments for description and recognition of moving Shapes'. Proc. British Machine Vision Conf. (BMVC01), 2001, pp. 705–714

20 Li, Sh., Lee, M.C., Pun, C.M.: 'Complex Zernike moments features for shape-based image retrieval', *IEEE Trans. System, Man, Cybern. – Part A: Syst. Humans*, 2009, **39**, (1), pp. 227–237

21 Li, X., Song, A.: 'A new edge detection method using Gaussian-Zernike moment operator'. Proc. Second Int. Asia Conf. Informatics in Control, Automation and Robotics (CAR), 2010, pp. 276–279

22 Papakostas, G.A., Karras, D.A., Mertzios, B.G.: 'Image coding using a wavelet based Zernike moments compression technique'. Proc. 14th Int. Conf. Digital Signal Processing, 2002, pp. 517–520

23 Teh, C.H., Chin, R.T.: 'On image analysis by the methods of moments', *IEEE Trans. Pattern Anal. Mach. Intell.*, 1988, **10**, (4), pp. 496–513

24 Lin, H., Si, J., Abousleman, G.P.: 'Orthogonal rotation-invariant moments for digital image processing', *IEEE Trans. Image Process.*, 2007, **17**, (3), pp. 272–282

25 Xin, Y., Pawlak, M., Liao, S.: 'Accurate computation of Zernike moments in polar coordinates', *IEEE Trans. Image Process.*, 2007, **16**, (2), pp. 581–587

26 Kim, W.Y., Kim, Y.S.: 'A region-based shape descriptor using Zernike moments', *Signal Process. Image Commun.*, 2000, **16**, (1), pp. 95–102

27 Hwang, S.K., Billinghurst, M., Kim, W.Y.: 'Local descriptor by Zernike moments for real-time keypoint matching'. Proc. IEEE Conf. Image Signal Processing, 2008, pp. 781–785

28 Schaefer, G., Stich, M.: 'UCID: an uncompressed color image database'. Proc. SPIE: Storage and Retrieval Methods and Applications for Multimedia, 2004, pp. 472–480

29 Gonzalez, R.C., Woods, R.E.: 'Digital image processing' (Pearson Education, 2008, 3rd edn.)

30 Peng, H., Long, F., Ding, C.: 'Feature selection based on mutual information: criteria of max-dependency, max-relevance, and min-redundancy', *IEEE Trans. Pattern Anal. Mach. Intell.*, 2005, **27**, (8), pp. 1226–1238

doi: 10.1049/iet-ipr.2012.0717