JOURNAL OF FORENSIC SCIENCES

J Forensic Sci, March 2015, Vol. 60, No. 2
doi: 10.1111/1556-4029.12680
Available online at: onlinelibrary.wiley.com

TECHNICAL NOTE

DIGITAL & MULTIMEDIA SCIENCES

Fei Peng,[1] *Ph.D.; Jiao-ting Li,*[1] *M.S.; and Min Long,*[2] *Ph.D.*

Identification of Natural Images and Computer-Generated Graphics Based on Statistical and Textural Features*

ABSTRACT: To discriminate the acquisition pipelines of digital images, a novel scheme for the identification of natural images and computer-generated graphics is proposed based on statistical and textural features. First, the differences between them are investigated from the view of statistics and texture, and 31 dimensions of feature are acquired for identification. Then, LIBSVM is used for the classification. Finally, the experimental results are presented. The results show that it can achieve an identification accuracy of 97.89% for computer-generated graphics, and an identification accuracy of 97.75% for natural images. The analyses also demonstrate the proposed method has excellent performance, compared with some existing methods based only on statistical features or other features. The method has a great potential to be implemented for the identification of natural images and computer-generated graphics.

KEYWORDS: forensic science, image source identification, natural images, computer-generated graphics, multifractal dimension, lacunarity analysis

With the boom of digital technologies, digital equipments such as digital cameras and digital scanners are widely used in our daily life. At the same time, more and more easy-to-learn but powerful image processing software are used to retouch images. As the software brings convenience to us, it also confuses us in the identification of the authenticity of images in some occasions, such as in news report, scientific research, insurance, and court evidence. Therefore, the research on the forensics of digital images is significant.

Currently, digital image forensics can be classified into two categories: active forensics and passive forensics. Active forensics mainly includes digital signature and digital watermarking. They usually need extra information, which limits its application; Passive forensics can identify the authenticity or the source of an image only using its own characteristics, and no extra information is needed (1), so it is more desirable in real applications. According to different purposes of the forensics, passive forensics in digital image can be further classified into image source forensics (2,3), image tamper forensics (4), and steganalysis.

Digital image acquisition pipeline forensics belongs to image source forensics (5). At present, digital image acquisition pipelines mainly include digital camera, digital scanner, and 3D modeling software. Recently, as much progress has been made in 3D modeling techniques, it is very likely to create a computer-generated graphics, which is similar to a natural image (their visual difference is very tiny). Therefore, how to identify natural images and computer-generated graphics is becoming an urgent problem to solve. This study aims to propose a more effective scheme for the identification of natural images and computer-generated graphics, where statistical and textural features are both considered for the forensics.

The rest of the paper is organized as follows: the related works are introduced in the next section; Feature analysis and selection are represented in the section which follows the next section. Continuing, the forensics scheme is described in the other section which follows it. Experimental results and analyses are discussed in the penultimate section. Finally, some conclusions are drawn in the last Section.

Related Works

Currently, some works have been conducted to identify the natural images and the computer-generated graphics. The existed methods can be classified into three categories according to the features used for identification (1): the methods based on the statistical features of the image (6,7), the methods based on the textural features of the image (8–10), and the methods based on the physical characteristics left by the image acquisition equipment (2–4,11–15).

As for the research based on statistical features, Lyu and Farid (6) proposed a statistical model based on the first-order and higher order wavelet statistical characteristics. Three-level wavelet decomposition is first applied to each color channel of an image, then mean, variance, skewness, and kurtosis of the coefficient histograms of each wavelet sub-band and prediction error

[1]School of Computer Science and Engineering, Hunan University, 410082 ChangSha, China.
[2]College of Computer and Communication Engineering, ChangSha University of Science and Technology, 410114 Changsha, China.

*Supported by National Natural Science Foundation of China (Grant No. 61370225, 61070195, 61001004), Hunan Provincial Natural Science Foundation of China (Grant No. 12JJA006), the Science and Research Plan of Hunan Province (Grant No. 2014FJ4161) and the Education Department Foundation of Hunan Province (Grant No. 11B002).

Received 6 Mar. 2013; and in revised form 16 Dec. 2013; accepted 11 Mar. 2014.

sub-band are selected as features. A Support Vector Machine (SVM) classifier is employed to 216 dimensions of features, with a classification accuracy of 98.8% for natural images and 66.8% for computer-generated graphics. Although the dimension of the features is large, the accuracy for computer-generated graphics is still limited, which indicates that the statistical features cannot sufficiently reflect the differences between these two pipelines alone. Guo et al.(7) proposed a method to identify real images and computer-generated graphics using the mean values of the wavelet coefficients in the prediction error image in HSV (Hue, Saturation, Value) color space. This method is based on the differences of the statistical features between natural images and computer-generated graphics.

As for the research based on textural characteristics, Ng et al.(8) proposed an image geometry model for the identification. It is found that there exist essential differences between the geometric object models of natural images and computer-generated graphics. The local fractal dimension and the image fragments are used to describe the visual features. One hundred Ninety-two dimensions of features are used for identification, and an average identification accuracy of 83.5% is obtained. However, the identification accuracy for computer-generated graphics of nature scenes is only 49.2%. Pan et al. (9) proposed an identification method based on fractal geometry. The difference in color perception between natural images and computer-generated graphics is represented by fractal dimension. Thirty dimensions of features are used, and an average accuracy of 91.2% is obtained.

As for the research based on physical characteristics left by the image acquisition equipment, Fridrich et al. applied the principle of digital watermarking technology for digital image forensics, regarding the reference pattern noise as a spread-spectrum watermark (11–13). The reference pattern noise of each camera is acquired by averaging the noise obtained from multiple images generated by each camera. Correlation between the pattern noise of an image and the corresponding reference pattern noise is used to distinguish the cameras models. After that, the residual pattern noise is used by Khanna et al. for the identification of natural images, scanning pictures, and computer-generated graphics (14,15). Fifteen features are used and an average identification accuracy of 85.9% is obtained. Dirik et al. (4) considered all interpolated and noninterpolated pixels, respectively, as a whole, and employed the noise ratio of these two pixel groups to distinguish natural images and computer-generated graphics. It is expected that all interpolated and noninterpolated pixels have the same properties. In practice, due to the influence of the interpolation algorithm, the content of the image and the nonlinear operation, it is difficult to ensure all pixels in these two groups have the same characteristics. Hence, if the pixels are simply divided into two groups and considered as a whole, the overall characteristics will be interfered by the interaction among pixels.

Besides the aforementioned three categories, Peng and Liu (16) proposed to identify natural images and computer-generated graphics based on hybrid features including statistical, textural, and physical characteristics. The mean, variance, kurtosis, skewness, and median of the histograms of grayscale image in the spatial and wavelet domain are selected as statistical features, the fractal dimensions of grayscale image and wavelet sub-bands are extracted as visual features, and the mean, variance, skewness, kurtosis and fractal dimensions of the enhanced PRNU are extracted as physical features. With a total of 48 dimensions of features, an average classification accuracy of 94.29% is achieved. However, the capability of discriminating natural images and computer-generated graphics is different.

From the above analysis, although much progress has been made in the identification of natural images and computer-generated graphics, prior works still have the following drawbacks: (i) Even if the dimension of features is large, the improvement of the identification accuracy is still limited; (ii) Most methods only acquire one aspect of the characteristics, which cannot fully reflect the differences between natural images and computer-generated graphics. As the contents of the images are abundant, if only a particular aspect of characteristics is considered, the amount of information obtained will be relatively limited.

To overcome the above limitations, a novel method based on the statistical and textural features is proposed to identify natural images and computer-generated graphics in this study.

Feature Selection

In this study, to describe the differences between natural images and computer-generated graphics, statistical and textural features are both investigated.

Statistical Feature

Statistical features are inherent to the content of images. Different images generally have different statistical properties. Due to the difference of imaging models and light transmissions between natural images and computer-generated graphics, there exist essential differences in their content. Thus, the statistical characteristics are important for the identification of natural images and computer-generated graphics. In previous studies of statistical models, the statistical features are generally extracted from the RGB (Red, Green, and Blue) color channels, and the number of dimensions is large (6). However, the analysis finds that the distribution of the histogram of the grayscale image is very similar to those of the RGB color channels, which is shown in Fig. 1. Hence, to reduce the feature dimension, only the statistical characteristics of the grayscale image are considered in this study.

In this study, some statistical features are obtained for the detection of natural images and computer-generated graphics. First, the color image I is transformed into a grayscale image I_g and then I_g is partitioned into $N(I)$ sub-blocks with a size of $bs \times bs$. The total number $N(I)$ is

$$N(I) = \lfloor M/bs \rfloor \times \lfloor N/bs \rfloor \tag{1}$$

Relative Frequency—As natural images are generally obtained by cameras, they have a different imaging principle from that of computer-generated graphics. The gray series of the natural images are generally more abundant than those of computer-generated graphics. As the computer-generated graphics are obtained by the 3D image software, it cannot completely simulate the natural scene. Here, the relative frequency is used to represent the characteristics of a pixel value. The calculation of the relative frequency of every pixel in each sub-block of the image I_g is:

$$F^k_{i,j} = P^k_{i,j}/(bs \times bs) 0 < k \le N(I), 0 < i \le bs, 0 < j \le bs \tag{2}$$

where k represents the k^{th} sub-block, $P^k_{i,j}$ represents the number of each pixel value that appears in its neighboring window with a size of 5×5, and $F^k_{i,j}$ is the relative frequency which represents the frequency of the pixel value in a sub-block. Hence, the relative frequency of all pixels in the k^{th} sub-block can be expressed as:

$$F^k = \{F^k_{i,j} | 0 < k \le N(I), 0 < i \le bs, 0 < j \le bs\} \tag{3}$$

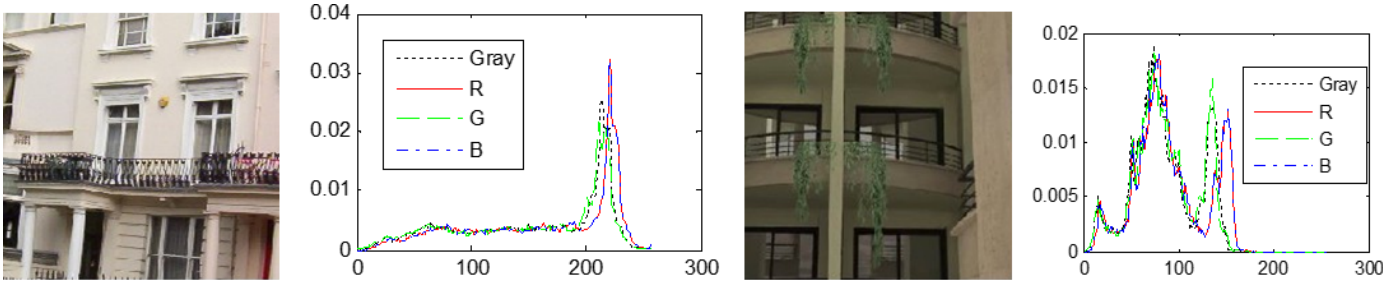

FIG. 1—*Histograms of RGB image and grayscale image. (a) A natural image; (b) Histograms of (a) and its grays-scale counterpart. (c) A computer-generated graphics. (d) Histograms of (c) and its gray-scale counterpart.*

Mean of Relative Frequency—Mean is an important statistical characteristic which can objectively reflect the overall distribution of a series of data. For a digital image, the relative frequency of every pixel represents its overall statistical characteristics, and its mean can describe the degree of light and shade of the image. Here, mean of relative frequency (MRF) is used to describe the average of the relative frequency in an image. MRF is calculated as

$$MRF = \frac{1}{N(I)} \sum_{k=1}^{N(I)} M^k \tag{4}$$

where M^k is the mean of the set F^k.

Variance of Relative Frequency—Variance reflects the degree of deviation from the mean value. For a digital image, the variance can describe the degree of information and the energy of the image, the greater the variance, the more the energy and information is. Hence, variance of relative frequency (VRF) is used to describe the statistical information in the relative frequency of the image. It is calculated as

$$VRF = \frac{1}{N(I)} \sum_{k=1}^{N(I)} V^k \tag{5}$$

where V^k is the variance of the set F^k.

To describe the differences of statistical features between natural images and computer-generated graphics, M^k of all pixels in the k^{th} sub-block of the grayscale image in the neighboring window with a size of 5 × 5 and bs = 16 is calculated and the statistical array $P = [M^1, M^2, \ldots M^k \ldots, M^{N(I)}]$ is then obtained. Through counting M^k of all images in the image data set, it is found that almost all values of M^k are distributed in [0, 0.1]. Hence, 0.01 is taken as an interval of the range [0, 0.1] and count the frequency of each data in P, thus the histogram is obtained, as seen in Fig. 2

From Fig. 2, it can be found that the value of the statistics histogram of natural images is greater than computer-generated graphics in each group. For this reason, these statistical features can be implemented for the identification of the natural images and computer-generated graphics.

Textural Features

There are also differences in texture between natural images and computer-generated graphics. A simple way to describe the texture is the statistical moments of an image's histogram. However, only histogram is not enough for the identification. Other textural features need to be considered. In this study, lacunarity of relative frequency (LRF), smoothness of relative frequency (SRF), entropy of relative frequency (ERF), consistency of relative frequency (CRF), and multifractal dimension are used as textural features.

Lacunarity is thought as a measure of the "gap" or "hole" of a geometric structure. It is defined as the nondimensional ratio of the second and first moments of mass distribution and obtained using the gliding-box algorithm (17). It also provides a parsimonious analysis for the overall fraction of a map or transect covered by the attribute of interest, the presence of self-similarity, the presence and scale of randomness, and the existence of hierarchical structure. Given an image I with a size of $M \times N$, its lacunarity $L(I)$ is defined as

$$L(I) = \frac{\sigma^2 + \mu^2}{\mu^2} \tag{6}$$

where μ is the mean of the pixels' gray scale in the image I, and σ^2 represents the variance.

Lacunarity analysis suggests an ordering of irregularity in the images. It has been widely used in many occasions, such as diagnosis in clinical application (18), measurement of space-filling capacity and heterogeneity (19), quantitation of the distribution of spaces between rock fractures (20). In this study, lacunarity texture is used as one aspect of feature to identify natural images and computer-generated graphics.

Based on the analysis about the lacunarity characteristics of images, some statistical measures such as lacunarity of relative frequency (LRF), smoothness of relative frequency (SRF), entropy of relative frequency (ERF), and consistency of relative frequency (CRF) are defined as follows.

Lacunarity of Relative Frequency—Lacunarity of Relative Frequency represents an ordering of irregularity and measurement of space-filling capacity and heterogeneity in the relative frequency of images. As natural images have complex texture and hierarchical structure features, which are different from computer-generated graphics, LRF is used to describe the texture information and the existence of hierarchical structure in the relative frequency of image. LRF is defined as:

$$LRF = \frac{1}{N(I)} \sum_{k=1}^{N(I)} \frac{(M^k)^2 + V^k}{(M^k)^2} \tag{7}$$

Smoothness of Relative Frequency—Smoothness is an important characteristic of the texture of an image. It represents the

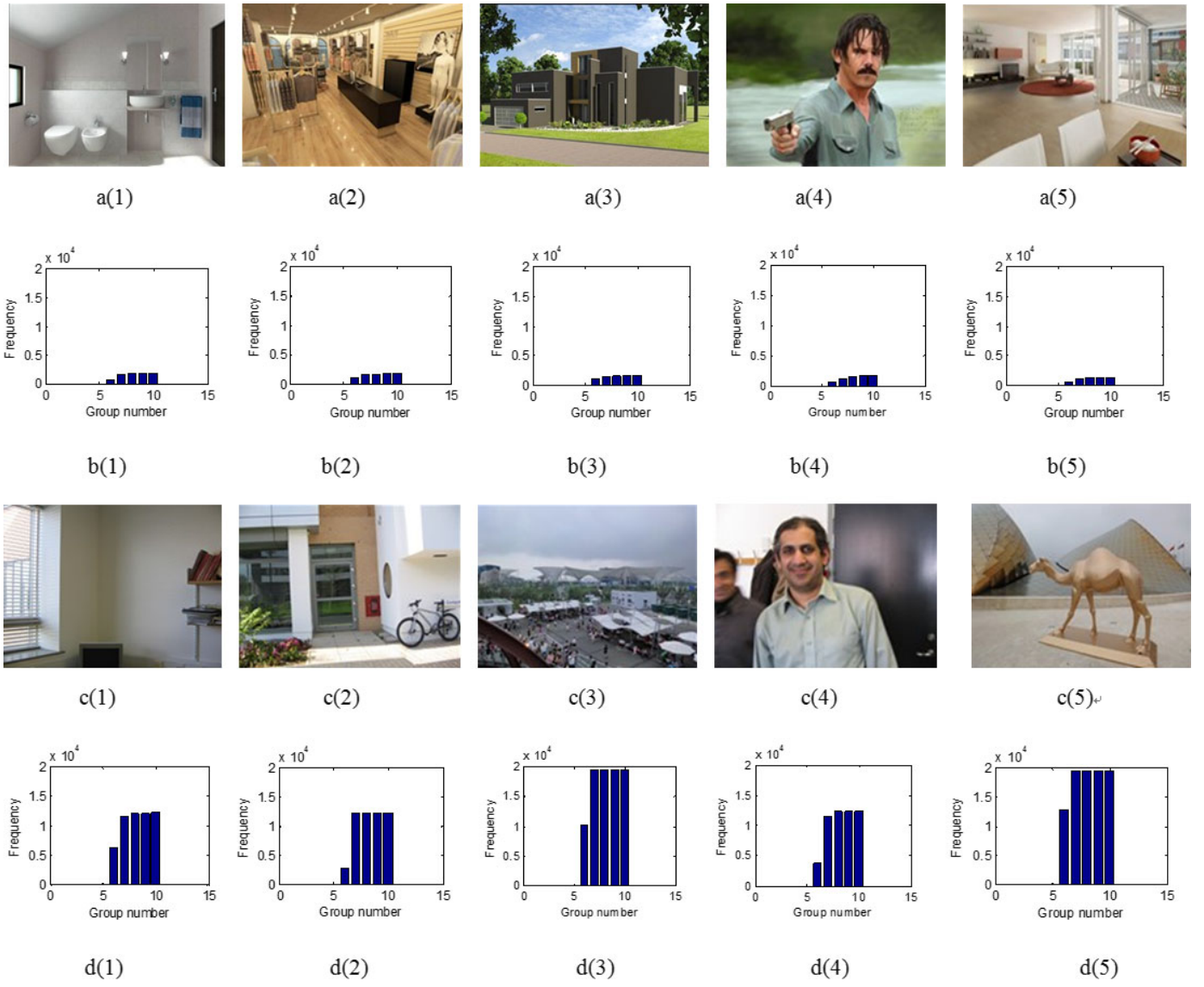

FIG. 2—*a(1)–a(5) some samples of computer-generated graphics (CG), b(1)–b(5) The corresponding* RFM *histograms of them, c(1)–c(5) some samples of Natural Images (NI), d(1)–d(5) The corresponding RFM histograms of them.*

average brightness of an image. SRF is a quantitation of the distribution of texture and measurement of smoothness in the relative frequency of an image. There are differences in smoothness between natural images and computer-generated graphics because of their different acquisition mechanisms. SRF is defined as:

$$SRF = \frac{1}{N(I)} \sum_{k=1}^{N(I)} \left(1 - \frac{1}{1 + V^k}\right) \tag{8}$$

Entropy of Relative Frequency—Entropy is a concept of energy. Image entropy describes the average amount of information of an image, which reflects the overview of an image and the image visual features. Natural images are obtained in a natural way, while computer-generated graphics rely on the individuals' imagination. Hence, the natural images are more complicated than computer-generated graphics. For this case, image entropy is an important characteristic for the identification of natural images and computer-generated graphics. ERF is calculated as:

$$ERF = -\frac{1}{N(I)} \sum_{k=1}^{N(I)} \sum_{i=1}^{bs} \sum_{j=1}^{bs} \left(F_{i,j}^k \times \log_2(F_{i,j}^k)\right) \tag{9}$$

Consistency of Relative Frequency—There exists the abundant information in the edge of an image. Image edges usually show the following characteristics: gray mutation, different regional boundary, and directivity. Consistency is a quantity to measure the difference in the image edge. According to the generation mechanism of natural images and computer-generated graphics, there exist obvious textural characteristics in natural images. Therefore, image consistency can be used for the representation of the difference between natural images and computer-generated graphics. CRF is a measurement of consistency in the relative frequency of image, and it is defined as:

$$CRF = \frac{1}{N(I)} \sum_{k=1}^{N(I)} \sum_{i=1}^{bs} \sum_{j=1}^{bs} (F_{i,j}^{k})^2 \tag{10}$$

Multifractal Dimension—Fractal can essentially describe the complexity and self-similarity of an object. Fractal dimension is generally used to represent the textural difference between images. However, a single fractal dimension cannot fully depict an image's textural characteristics, because some images with the same fractal dimension may have great differences in the visual effect. To provide a thorough description of fractal, multifractal is introduced to describe the characteristics of different fractal sets (21).

As known to us, computer-generated graphics is obtained from 3D models after rendering operations. Its texture is generally simpler than that of natural images. Here, generalized multifractal dimension is defined to represent the difference of the textural characteristics between computer-generated graphics and natural images.

Given an image I, it is partitioned into sub-blocks with a size of $r \times r$, each sub-block is represented as C_i, and i is its index. A probability mass function, $\mu(C_i)$, is defined as:

$$\mu(C_i) = \frac{\sum_{C_i} I(m,n)}{\sum_{I} I(m,n)} \tag{11}$$

Its q order moment can be represented as:

$$S_r(q) = \sum_i \mu(C_i)^q \tag{12}$$

For a constant r, the value of $S_r(q)$ decreases as q increases. As $r \to 0$, $T(q)$ is obtained as:

$$T(q) = \lim_{r \to 0} \frac{\log S_r(q)}{\log(1/r)} \tag{13}$$

For different r, the value of $T(q)$ can be obtained using the least squares fitting method. Finally, a multifractal D_q is calculated as:

$$D_q = \frac{T(q)}{q-1} \tag{14}$$

Actually, *LRF*, *MRF*, and *VRF* are all calculated in the wavelet domain in the proposed scheme. This is because that the wavelet transformation is a kind of time–frequency transformations and can represent the digital images in a more redundant form (22).

Description of the Identification Scheme

Given a color image I, the extraction of identification features is illustrated in Fig. 3, and the details of the extraction can be described in the following:

Step 1. Transfer I into a grayscale image I' and obtain the relative frequency F^k of all pixels according to the method described in Section "Statistical Feature", then the *SRF, ERF,* and *CRF* are calculated as the statistical features of I'.

Step 2. Decompose I' with the two-level "Daubechies-8 wavelet" wavelet transform, and all 7 sub-bands of each component are selected for feature extraction. The wavelet coefficients of each sub-band are transformed into integers with floor operation, then *LRF*, *MRF*, and *VRF* are calculated according to the method described in Section "Feature Selection".

Step 3. Multifractal dimensions $T(q)$ (q = 2, 4, 6, 8, 10, 12, 14) are calculated from I' according to the method described in Section "Relative Frequency". The above 31 dimensions of feature are used as the input of SVM classifier, which is a supervised learning model with associated learning algorithms that analyze data and recognize patterns (23). It is used to determinate whether the image I is a natural image or a computer-generated graphics.

Step 4. The image data sets are divided into a training set and a testing set. The training set is used to train the SVM classifier to obtain the best identification model, while the testing set is used to predict the performance of the identification model. In a SVM classifier, the class label of the natural image is 1 and the class label of the computer-generated graphics is −1. The identification process includes the training process and the testing process, which are shown in Fig. 4(*a*) and (*b*), respectively.

Training Process

Step 1. Input all images in the training data set;

Step 2. For each image, extract statistical and textural features according to the above feature extraction process from *Step* 1 to *Step* 3, with a total of 31 dimensions;

Step 3. Use the features of the training data set to train the SVM classifier, with a kernel function for the RBF and a fivefold cross-validation to find the optimal parameters and construct a SVM model.

Testing Process

Step 1. Input all images in the testing data set;

Step 2. For each image, extract statistical and textural features according to the above feature extraction process from *Step* 1 to *Step* 3, with a total of 31 dimensions;

Step 3. Input their features to the SVM classifier and obtain their corresponding classification results for either natural image or computer-generated graphics.

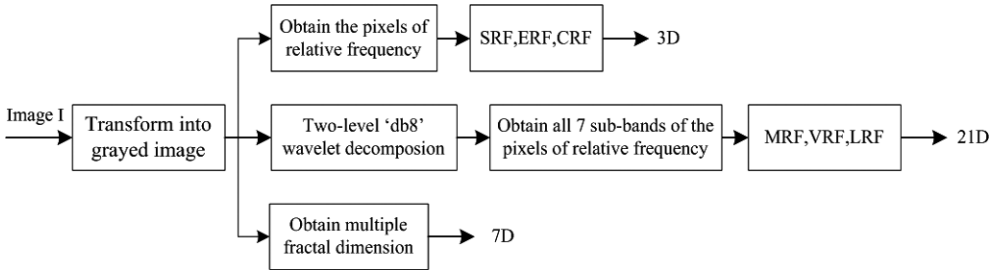

FIG. 3—*Feature extraction process.*

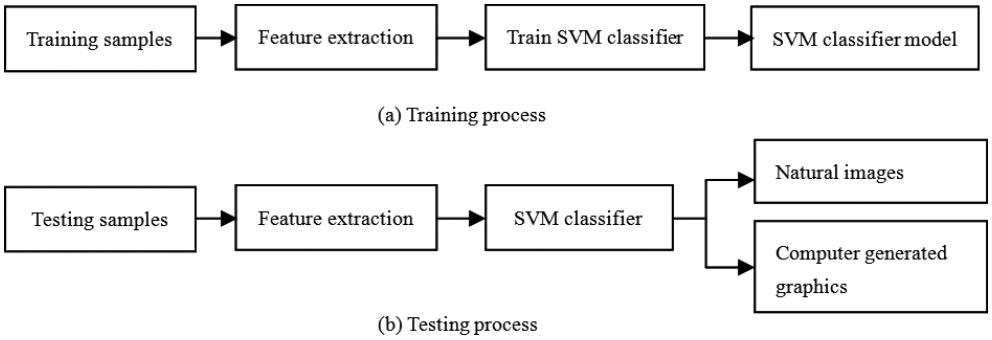

FIG. 4—*The diagram of the identification process.*

Experimental Results and Analysis

In the experiments, an image database containing 2400 JPEG images is used, where the number of natural images and computer-generated graphics are both 1200, respectively. The camera models and brands are listed in Table 1. The acquisition time of natural images includes day and night, the shooting temperature is different, and a variety of outdoor and indoor scenes are included. High photorealistic computer-generated graphics come from the image data sets at Columbia University and at some 3D image websites listed in Table 2. The image processing software involve MAYA, 3D MAX, Poser, Softimage XSI and so on. Here, only JPEG images are used in the experiments due to their wide application in daily lives.

Experimental Results

In the experiments, Matlab2010, Python2.5, and gnuplot4.2 are used as the simulation tools. An IBM compatible PC with 1 GB RAM and 2.20 GHz CPU is used for experiment, and LIBSVM (24) is used as the classifier for the identification. To ensure the objectivity of the experiments, the training set and testing set are randomly chosen by subset.py. Here, 800 images randomly chosen from each type are used for training and the rest are used for testing, and LIBSVM is applied to these two types to identify their source acquisition pipelines. All experimental results are the average results of 10 randomly repeated experimental runs.

To analyze the sensitivity of the proposed method to the selection of database, another image database is setup. The original image database is named as Database 1, and the new image database is named as Database 2. After the experiments, an average classification accuracy of 97.82% is achieved for Database 1 using the SVM classifier, where the classification for computer-generated graphics is 97.89% and that for natural images is 97.75%. It can be found that using the SVM classifier, the classification for Database 2 is very similar to that for Database 1. The results are shown in Table 3.

As seen in Table 3, the identification performance for Database 1 and Database 2 is very similar, which indicates that the proposed scheme is not sensitive to the image database.

Analysis of the Sensitivity to Classifiers

To evaluate the sensitivity of the proposed scheme to classifiers, another classifier Linear Discriminant Analysis (LDA) is used in the experiments with the same image databases. The experimental results are shown in Table 4.

Comparing the results in Table 4 with those in Table 3, SVM outperforms LDA in classification accuracy. It is because LDA needs linear classification surface, and the features are often projected to 1 dimension in the classification. What is more, LDA is not suitable for non-Gaussian distribution samples. No matter how it is projected in line, it is difficult to completely separate them. While LIBSVM is a nonlinear SVM classifier, it can map a lower dimension space into a high dimensional space and

TABLE 1—*Cameras used in experiments.*

No.	Cameral Model	Resolution	No.	Cameral Model	Resolution
1	Canon IXUS 850	600 × 800	8	Sony T70	2048 × 1536
2	Canon IXUS 850	2048 × 1536	9	Sony T700	2048 × 1536
3	Canon IXY 500	2048 × 1536	10	Sony WX1	2592 × 1944
4	Canon IXY 800	2816 × 2112	11	BenQ C1020	3648 × 2736
5	Canon PowerShot A40	2048 × 1536	12	BenQ T850	3264 × 2448
6	FujiFilm S602	2048 × 1536	13	Olympus C730UZ	2048 × 1536
7	FujiFilm A920	2048 × 1536	14	Kodak CX4230	816 × 616

TABLE 2—Computer-generated graphics web sites used in experiments.

The URLs of Computer-generated graphics webs
http://www.archimodel.com, http://www.accurender.com, http://gallery.mcneel.com, http://www.npowersoftware.com, http://www.marlinstudios.com, http://www.realsoft.fi, http://www.kotiposti.net, http://www.mentalimages.com, http://www.artlantis.com, http://www.3dshop.com, http://www.flamingo3d.com, http://www.evs3d.com, http://www.maxoncomputer.com, http://raph.com, http://www.xfrogdownloads.com

TABLE 3—*Experimental results with LIBSVM.*

	Database 1		Database 2	
	NI	CG	NI	CG
NI %	97.89	2.11	95.65	4.35
CG %	2.25	97.75	2.76	97.24

NI, Natural Images; CG, computer-generated graphics

TABLE 4—*The experimental results.*

	Database 1		Database 2	
	NI	CG	NI	CG
NI %	85.50	14.50	87.51	12.49
CG %	10.52	89.48	11.50	88.50

NI, Natural Images; CG, computer-generated graphics

make them linear separable. So, it is more flexible for image classification. For this reason, the remaining experiments are all carried out using LIBSVM.

Analysis of the Effectiveness of the Features

To analyze the effectiveness and the contribution of each type of features to the identification results, the experiments are carried out regarding different kinds of features, respectively. The results are shown in Table 5.

As seen in Table 5, it can be found that the identification accuracy for 14 dimensions of statistical features is 94.38% for natural images and 93.26% for computer-generated graphics. The identification accuracy for 17 dimensions of textural features is 95.63% for natural images and 90.55% for computer-generated graphics. The above results indicate that the selected two types of features are both effective.

To analyze the influence of the value of q on the identification, the experiments are also conducted to validate the selection of the values of q. The results are shown in Fig. 1(a) and (c). The relation between D_q and q is shown in Fig. 5, where the value of D_q for Fig. 1(a) and (c) is both almost zero when q takes negative values. There also exists the obvious difference when q takes the positive values. As seen in Fig. 5, with the increase of the value of q, D_q is becoming a constant value. Thus, q = (2, 4, 6, 8, 10, 12, 14) are chosen for the calculation of multifractal dimension.

Analysis of the Performance

To analyze the performance of the proposed scheme, the experiments are also conducted to compare the methods in (4,6,7,16). The results are shown in Table 6.

As seen in Table 6, the proposed method can achieve a detection accuracy of 97.75% for computer-generated graphics, a

TABLE 5—*The experimental results of features group.*

	Statistical features (14D)	Textural features (17D)
NI %	94.38	95.63
CG %	93.26	90.55
AC %	93.82	93.09

NI, Natural Images; CG, computer-generated graphics; AC, average detection accuracy.

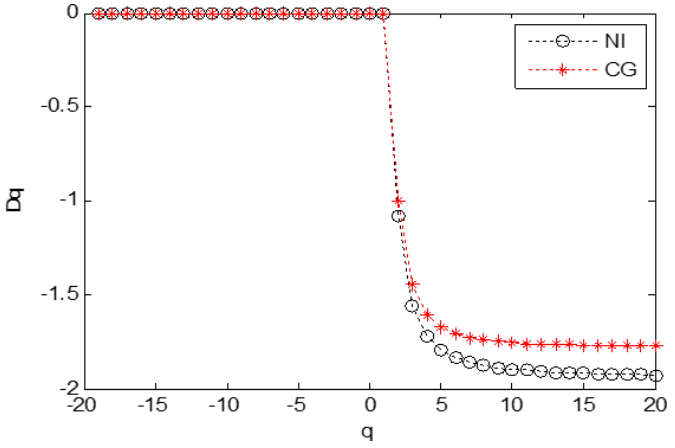

FIG. 5—*The relation between* Dq *and* q.

detection accuracy of 97.89% for natural images, which outperforms other methods in (4,6,7,16). Besides, it can be found that the capability of discriminating natural images and computer-generated graphics is nearly the same, which also illustrated the good performance of the proposed method.

To further analyze the performance of the methods based on statistical features (6,7), CFA interpolation features (4), hybrid features (16), and the proposed method, the ROC(receiver operating characteristic) curves of these methods are obtained, as shown in Fig. 6.

From Fig. 6, it can be found that the AUC (Area Under Curve) of the proposed method is about 0.99, and it has a better performance than others, which indicates the good performance of the proposed scheme.

Analysis of the Ratio between the Training set and the Testing Set

To analyze the influence of the ratio between the training set and the testing set, the experiments are conducted with different ratios. The results are shown in Table 7.

From Table 7, it can be found that the experimental results have little fluctuation when the ratio varies. The worst detection accuracy is 93.45% when the ratio between training samples and testing samples is 1:5. It is possible that only 400 sample images are used for training, which is not sufficient for the construction of an accurate classification model. When the ratio is 2:1, it can obtain the best identification performance. Hence, this ratio is recommended to be used for the classification.

Analysis of the Robustness

To analyze the robustness of the proposed scheme, nine different image operations, including JPEG recompression with quality factors (QF) of 90, 70, and 50, adding noise with SNRs of 60, 40, 20, resizing with scale factors (SF) of 1/4, 1/3, and ½, are applied to all testing images. The experiments are carried out to these images with the proposed identification scheme. The results are shown in Table 8.

From Table 8, it can be seen that the detection accuracies are changed very little when they undergo the operations of JPEG compression and adding noise, which indicates its good ability of resisting JPEG compression and adding noise. However, the detection accuracies are greatly changed when they undergone

TABLE 6—*The experimental results of textural features.*

	Literature [6] (216D)	Literature [7] (39)	Literature [4] (1D)	Literature [16] (48D)	The proposed method (31D)
NI %	90.19	81.12	72.05	91.28	**97.89**
CG %	88.97	84.54	72.59	97.3	**97.75**
AC %	89.58	82.83	72.32	94.29	**97.82**
AUC	0.93	0.87	0.79	0.98	**0.99**

NI, Natural Images; CG, computer-generated graphics; AUC, Area Under Curve; AC, average detection accuracy.

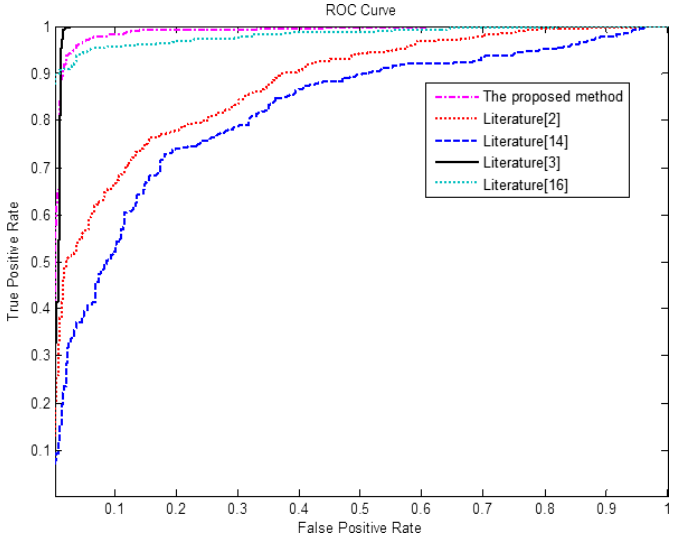

FIG. 6—*ROC curves of different methods.*

TABLE 7—*The experimental results on different samples distributions.*

Training set: testing set	1:5	1:2	1:1	2:1	5:1
NI %	93.12	95.51	95.43	**97.89**	96.68
CG %	93.86	95.63	96.83	**97.75**	96.98
AC %	93.49	95.57	96.13	**97.82**	96.83

NI, Natural Images; CG, computer-generated graphics; AC, average detection accuracy.

TABLE 8—*Analysis results of the robustness.*

Operations	NI %	CG %	AC %
JPEG QF90	97.85	97.51	97.68
JPEG QF70	97.82	97.48	97.65
JPEG QF50	96.87	92.23	94.55
SNR 60	97.15	96.97	97.06
SNR 40	97.17	97.07	97.12
SNR 20	97.83	97.55	97.69
SF 1/4	93.26	82.52	87.89
SF 1/3	85.15	75.25	80.20
SF 1/2	94.18	90.22	92.20

NI, Natural Images; CG, computer-generated graphics

resizing. The possible reason for this could be that some important information of images was lost during the resizing.

Conclusions

In this study, the statistical and textural features are both taken into consideration for the identification of natural images and computer-generated graphics. Based on the differences in their principles of the image generation process, the significant difference between natural images and computer graphics in statistical features and textural features are investigated and 31 dimensions of features are selected for forensics. The experimental results and their analyses show that the proposed scheme can achieve a good performance, and it is robust against JPEG compression and adding noise. Although the textural features have a good contribution to the classification of natural images, the identification accuracy for computer-generated graphics is still limited. To further improve the identification accuracy, our future work will concentrate on the selection of other effective features in the view of fractal geometry, such as porosity and multifractal spectrum of images.

Acknowledgments

The authors thank the anonymous reviewers for their kind suggestions and also thank Dr. L. He, who is with the Department of Computer Science, University of Warwick, for his kind proofreading of this manuscript.

Reference

1. Chen DM, Li JH, Wang SL, Li SH. Identifying computer generated and digital camera images using fractional lower order moments. Proceedings of the Fourth IEEE Conference on Industrial Electronics and Applications, 2009 (ICIEA 2009); 2009 May 25-27; Xi'an, People's Republic of China. Piscataway, NJ: Institute of Electrical and Electronics Engineers, 2009;230–5.
2. Li CT. Source camera identification using enhanced sensor pattern noise. IEEE Trans Inf Forensics Secur 2010;5(2):280–7.
3. Li CT, Li Y. Colour decoupled photo response non-uniformity for digital image forensics. IEEE Trans Circuits Syst Video Technol 2012;22(2):260–71.
4. Dirik AE, Memon N. Image tamper detection based on demosaicing artifacts. Proceedings of the 16th IEEE International Conference on Image Processing (ICIP); 2009 Nov 7-10; Cairo, Egypt. Piscataway, NJ: Institute of Electrical and Electronics Engineers, 2009;1497–500.
5. Zhou LN, Wang DM. Digital image forensic techniques. Beijing, China: Beijing University, 2008.
6. Farid H, Lyu S. How realistic is photorealistic? IEEE Trans Signal Process 2005;53(2):845–50.
7. Guo K, Wang R. An effective method for identifying natural images and computer graphics. J Comput Inform Syst 2010;6(10):3303–8.
8. Ng TT, Chang SF. Physics-motivated features for distinguishing photographic images and computer graphics. Proceedings of the 13th Annual ACM International Conference on Multimedia (MM '05); 2005 November 6–12; Singapore. New York, NY: ACM, 2005;239–48; http://www.ee.columbia.edu/ln/dvmm/publications/05/ng_acmmm05.pdf.
9. Pan F, Chen JB, Huang JW. Discriminating between photorealistic computer graphics and natural images using fractal geometry. Sci China Ser F: Info Sci 2009;52(2):329–37.
10. Yang GQ, Cui RY. Method for natural image discrimination based on texture feature. Comput Appl Res 2010;27(7):2784–5.
11. Lukáš J, Fridrich J, Goljan M. Digital camera identification from sensor pattern noise. IEEE Trans Inf Forensics Secur 2006;1(2):205–14.
12. Lukáš J, Fridrich J, Goljan M. Determining digital image origin using sensor imperfections. Proc. SPIE 5685, Image and Video Communications and Processing. Bellingham, WA: SPIE, 2005;249–60; doi: 10.1117/12.587105.

13. Lukáš J, Fridrich J, Goljan M. Detecting digital image forgeries using sensor pattern noise. Proc. SPIE 6072, Security, Steganography, and watermarking of Multimedia Contents VIII, 60720Y. Bellingham, WA: SPIE, 2006;362–72; doi: 10.1117/12.640109.
14. Khanna N, Chiu GTC. Forensic techniques for classifying scanner. Proceedings of the 2008 IEEE International Conference on Computer Generated and Digital Camera Images, Acoustics, Speech and Signal Processing (ICASSP 2008); 2008 March 31-April 4; Las Vegas, NV. Piscataway, NJ: Institute of Electrical and Electronics Engineers, 2008;1653–6.
15. Khannaa N, Aravind K, George M, Chiu TC, Allebach P, Delp EJ. Forensic classification of imaging sensor types. Proc. SPIE 6505, Security, Steganography and Watermarking of Multimedia Contents IX. Bellingham, WA: SPIE, 2007;6505; doi: 10.1117/12.705849.
16. Peng F, Liu J. Identification of natural images and computer generated graphics based on hybrid features. Int J Digit Crime Forensics 2012;4 (1):1–16.
17. Plotnick RE, Gardner RH, O'Neill RV. Lacunarity indices as measures of landscape texture. Landscape Ecol 1993;8(3):201–11.
18. Gilmore S, Hofmann-Wellenh R, Muir J, Soyer HP. Lacunarity analysis: a promising method for the automated assessment of melanocytic naevi and melanoma. PLoS ONE 2009;4(10):e7449.
19. Valous NA, Mendoza F, Sun DW, Allen P. Texture appearance characterization of pre-sliced pork ham images using fractal metrics: fourier analysis dimension and lacunarity. Food Res Int 2009;42(3):353–62.
20. Roy A, Perfect E, Dunne WM, Odling N, Kim JW. Lacunarity analysis of fracture networks: evidence for scale-dependent clustering. J Struct Geol 2010;32(10):1444–9.
21. Nilsson E. Multifractal-based image analysis with applications in medical imaging [M.S. thesis]. Umea, Sweden: Dept. Computing Science, Umea Univ, 2007.
22. Akansu AN, Serdijn WA, Selesnick IW. Wavelet transforms in signal processing: a review of emerging applications. Physical Commun 2010;3:1–18.
23. Corinna C, Vladimir V. Support-vector networks. Mach Learn 1995;20 (3):273–97.
24. Chang CC, Lin CJ. LIBSVM: a library for support vector machines. ACM Trans Intell Syst Technol 2011;2(3):27.

Additional information and reprint requests:
Fei Peng, Ph.D.
School of Information Science and Engineering
Hunan University
No. 352, Lushannan Road
Yuelu District, Changsha, Hunan Province
410082
China
E-mail: pengfei@hnu.edu.cn