Soft Comput
DOI 10.1007/s00500-016-2146-6

METHODOLOGIES AND APPLICATION

Detection of content-aware image resizing based on Benford's law

Guorui Sheng[1] · **Tao Li**[1] · **Qingtang Su**[1] · **Beijing Chen**[2] · **Yi Tang**[3]

© Springer-Verlag Berlin Heidelberg 2016

Abstract Content-aware image resizing is currently widely used because it maintains the original appearance of important objects to the greatest extent when the aspect ratio of an image changes during resizing. Content-aware image resizing techniques, such as seam carving, are also used for image forgery. A new Benford's law-based algorithm for detecting content-aware resized images is presented. The algorithm extracts features on the basis of the first digit distribution of the discrete cosine transform coefficients, which follow the standard Benford's law. We trained these features from both normal images and content-aware resized images using a support vector machine. The experimental results show that the proposed method can efficiently distinguish a content-aware resized image from a normal image, and its precision is better than that of existing methods, including those based on Markov features and others.

Communicated by V. Loia.

✉ Guorui Sheng
shengguorui@outlook.com

Tao Li
litao_888@sina.com

Qingtang Su
sdytsqt@163.com

Beijing Chen
bjchen1876@163.com

Yi Tang
ytang.bjs@139.com

[1] School of Information Science and Electrical Engineering, Ludong University, Yantai, China

[2] School of Computer and Software, Nanjing University of Information Science and Technology, Nanjing, China

[3] Department of Mathematics, Guangzhou University, Guangzhou, China

Keywords Content-aware image resizing · Image forensics · Image forgery · Seam carving · SVM · Benford's law

1 Introduction

With the rapid development of digital image processing techniques, an increasing number of powerful digital image processing software applications have appeared, such as Adobe Photoshop and GIMP. These applications make image processing convenient, but they also make it far easier to tamper with images and misuse them. When altered images are used in news stories, as illustrations for scientific discoveries, in insurance fraud, or as evidence in courts of law, social order can be disrupted. False images change the true face of the world and may negatively affect the justice system. Therefore, determining whether an image is authentic is important. The field of image forensics deals with various ways to detect image tampering.

Image forensic techniques can be divided into two categories: proactive forensics and passive forensics. In proactive forensics, special information, such as a watermark, is embedded in the original normal image. This special information can later be extracted and compared with the original embedded information to check the authenticity of an image. Watermarking (Weng et al. 2013; Liu and Tsai 2010) is the primary proactive forensic technique, which, together with corresponding cryptography techniques (Li et al. 2014; Li and Kim 2010; Li et al. 2011; Wang et al. 2015), and digital signature techniques (Li et al. 2007; Wang et al. 2013; Li et al. 2010; Li and Kim 2010), takes advantage of redundant information in digital images. The primary limitation of proactive forensics is that when handling a large number of digital images, people are usually unable to know in advance

Published online: 25 April 2016

which images should be checked. In contrast, passive forensic techniques enable the authenticity of an image to be judged in advance, without extracting any embedded information. The broad applications of passive forensics have motivated researchers to focus on this field.

Common image tampering methods include copy-move, resampling, and splicing. Copy–move involves copying a portion of the original image and then pasting it to another portion of the same image. The method is usually used to hide or duplicate specific content, such as an object or a person. To detect such copy–move alterations, corresponding forensic methods seek to determine the presence of homogeneous regions present in the image using characteristics in the frequency domain of the image (Fridrich et al. 2003; Amerini et al. 2011, 2013). Image resampling includes operations such as resizing, rotating, and scaling. These operations resample the original image and cause detectable periodic correlations. Image splicing involves combining two or more images into one. The corresponding forensic image detection methods mainly focus on analyzing image features, such as high-order moment and bicoherence spectrum (Shi et al. 2007; Fu et al. 2006).

A novel and efficient content-aware image resizing technique called seam carving or seam insertion (hereinafter referred to as "Seam-CI") was proposed by Avidan (Avidan and Shamir 2007). This method resizes the original image by changing the aspect ratio while minimizing the resulting distortion to the content of the main image. Seam-CI is effective and has been added to Adobe Photoshop CS4 as an optional function and to GIMP as a plugin. However, Seam-CI can also be used to tamper with images maliciously, such as by removing specific objects or people from the original image.

1.1 Related Works

Three studies that focus on Seam-CI have been conducted so far. Min Wu proposed a hash-based detecting method (Lu and Wu 2011) in a study that focused on the problem of seam carving estimation and tampering localization by using compact side information called a forensic hash. The forensic hash technique tries to bridge two related approaches, namely, robust image hashing and blind multimedia forensics, to address a broad scope of forensic questions efficiently and accurately. The modular design of the forensic hash method avoids the one-scheme-fits-all dilemma. In addition, its compact constructions can provide a robust estimation of geometric transforms, such as rotation and scaling. The proposed forensic hash construction can be extended to estimate seam carving and detect local tampering accurately; however, the method is considered to be a proactive forensic technique, which means it has limited application.

Fillion proposed a multi-feature fusion-based method (Fillion and Sharma 2010). The work suggests that Seam-CI detection is a problem of classifying a test image into either a seam-carved or non-seam-carved category. He adopted a pattern recognition approach in which a set of features is extracted from the test image and then used a support vector machine (SVM)-based classifier trained on a set of images to estimate in which of the two classes the test image lies. The work developed four sets of features for detection. Two of these feature sets were motivated by an intuition that seam-carving will affect image statistics. The fourth set of features was generated by applying an additional seam-carving operation to the image. The method has a 91 % classification accuracy.

Sarkar proposed a Markov feature-based method (Sarkar et al. 2009). They employed a machine learning-based framework to distinguish between seam-carved (or seam-inserted) and normal images. The 324-dimensional Markov feature consists of 2D difference histograms in the block-based discrete cosine transform (DCT) domain. First, they chose to work on JPEG 2-D arrays to formulate features for steganalysis. Difference JPEG 2-D arrays along the horizontal, vertical and diagonal axes, as shown in Eq. (1), were then used to generally enhance the changes caused by JPEG steganography. The Markov process was applied to model these difference JPEG 2-D arrays to utilize the second order statistics for steganalysis, as shown in Eq. (2). In addition to the utilization of difference JPEG 2-D arrays, a thresholding technique was developed that greatly reduced the dimensionality of transition probability matrixes, i.e., the dimensionality of feature vectors, thus making the computational complexity of the scheme manageable.

$$\begin{aligned} F_h(u,v) &= F(u,v) - F(u+1,v) \\ F_v(u,v) &= F(u,v) - F(u,v+1) \end{aligned} \tag{1}$$

$$\begin{aligned} P1_h &= \frac{\sum_{u=1}^{S_u-2}\sum_{v=1}^{S_v}\delta(F_h(u,v)=i,F_h(u+1,v)=j)}{\sum_{u=1}^{S_u-2}\sum_{v=1}^{S_v}\delta(F_h(u,v)=i)} \\ P1_v &= \frac{\sum_{u=1}^{S_u-1}\sum_{v=1}^{S_v-1}\delta(F_h(u,v)=i,F_h(u,v+1)=j)}{\sum_{u=1}^{S_u-1}\sum_{v=1}^{S_v-1}\delta(F_h(u,v)=i)} \\ P2_h &= \frac{\sum_{u=1}^{S_u-1}\sum_{v=1}^{S_v-1}\delta(F_v(u,v)=i,F_v(u+1,v)=j)}{\sum_{u=1}^{S_u-1}\sum_{v=1}^{S_v-1}\delta(F_v(u,v)=i)} \\ P2_v &= \frac{\sum_{u=1}^{S_u}\sum_{v=1}^{S_v-2}\delta(F_v(u,v)=i,F_v(u,v+1)=j)}{\sum_{u=1}^{S_u}\sum_{v=1}^{S_v-2}\delta(F_v(u,v)=i)} \end{aligned} \tag{2}$$

The feature achieved detection accuracies of 80 and 85 % for seam carving and seam insertion, respectively. For seam insertion, each new pixel that is introduced is a linear combination of its neighboring pixels. Seam insertions are detected on the basis of this linear relation. The method achieved a high detection accuracy of 94 %.

1.2 Motivation and contribution

Benford's law, also called significant digit law or the first digit law, is an empirical law that was first discovered by Newcomb (1881) in 1881 and then rediscovered by Benford Benford

(1938) in 1938. A statistical explanation was given by Hill (1995). The law illuminates that the probability distribution of the first digits in a set of natural numbers is logarithmic.

DCT is currently the most popular image compression scheme. In this scheme, source image samples are first grouped into non-overlapping, consecutive 8×8 blocks, and then each block is transformed by forward DCT into a set of 64 values referred to as DCT coefficients. The value located in the upper-left corner of the block is called the direct current (DC) coefficient, and the other 63 values are called alternate current (AC) coefficients. The probabilities of the first digits of DCT coefficients from AC modes will be altered if a tampering operation used Seam-CI. .

In this paper, we propose using the probabilities of the first digits of DCT coefficients from individual AC modes as features to help detect seam-carved images. Seam-carved images and normal images can be differentiated using a two-class classification strategy. Experiments performed on the Columbia image database and CASIA, which are the acknowledged image databases, showed that our proposed method obtained satisfactory detection accuracy.

1.3 Paper outline

The rest of this paper is organized as follows: Sect. 2 briefly introduces the Seam-CI algorithm. Section 3 illustrates how closely the DCT coefficients follow Benford's law and then presents the techniques used to distinguish a tampered image from a normal image. Section 4 provides the experimental results and a comparison with other existing methods. Conclusions are drawn in Sect. 5.

2 Brief introduction of Seam-CI

Seam-CI is a content-aware image resizing technique proposed by Avidan (Avidan and Shamir 2007).

Normally, image cropping and image scaling are two ways to change an image's size. Image cropping can reduce the size of the image by cropping the pixels on the boundary while the remaining part is exactly the same as the original image. Cropping may remove some important contents of the image, which is probably on the boundary area. Image scaling also ignore the contents of the image, change the image size by interpolation, which will cause the correlation change between adjacent pixels. The correlation is depended on resizing scale and specific interpolation algorithm. Corresponding forensic methods are mainly based on detecting different kinds of correlation between adjacent pixels.

Unlike image cropping and image scaling, Seam-CI resizes an image by finding the "seam", which is a curve composed of 8-connected pixels that pass through the image horizontally or vertically. A seam has only one pixel in a column (horizontal seam) or a row (vertical seam). Two neighborhood pixels in a seam are 8-connected. Following the strategy of finding the seam, the sum of the seam's pixel's energy is ensured to obtain the minimum value. The low energy helps the seam to avoid containing the details of the image to the greatest extent. On the basis of this strategy, deleting or duplicating seams can resize the original image while retaining the most important contents of the original image. Comparison among image cropping, image scaling, and Seam-CI is shown in Fig. 1. Image cropping cropped out a discarded part

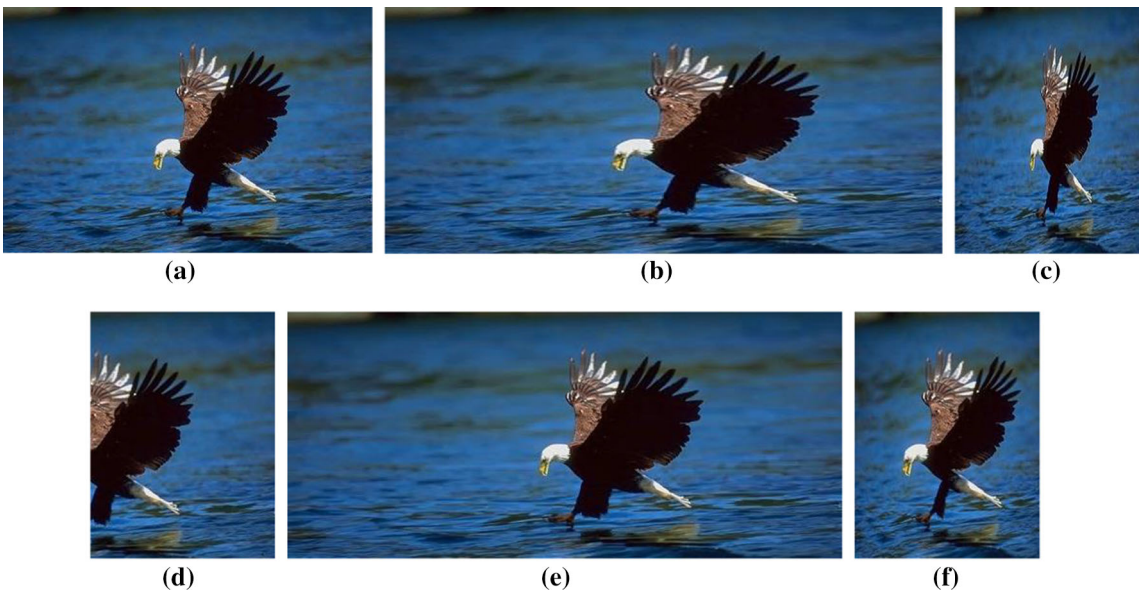

Fig. 1 Comparison of three image resizing methods: **a** original image **b** enlarged image by 50 % proportion using interpolation **c** reduced image by 50 % proportion using interpolation **d** reduced image by 50 % proportion using cropping **e** enlarged image by 50 % proportion using Seam-CI **f** reduced image by 50 % proportion using Seam-CI

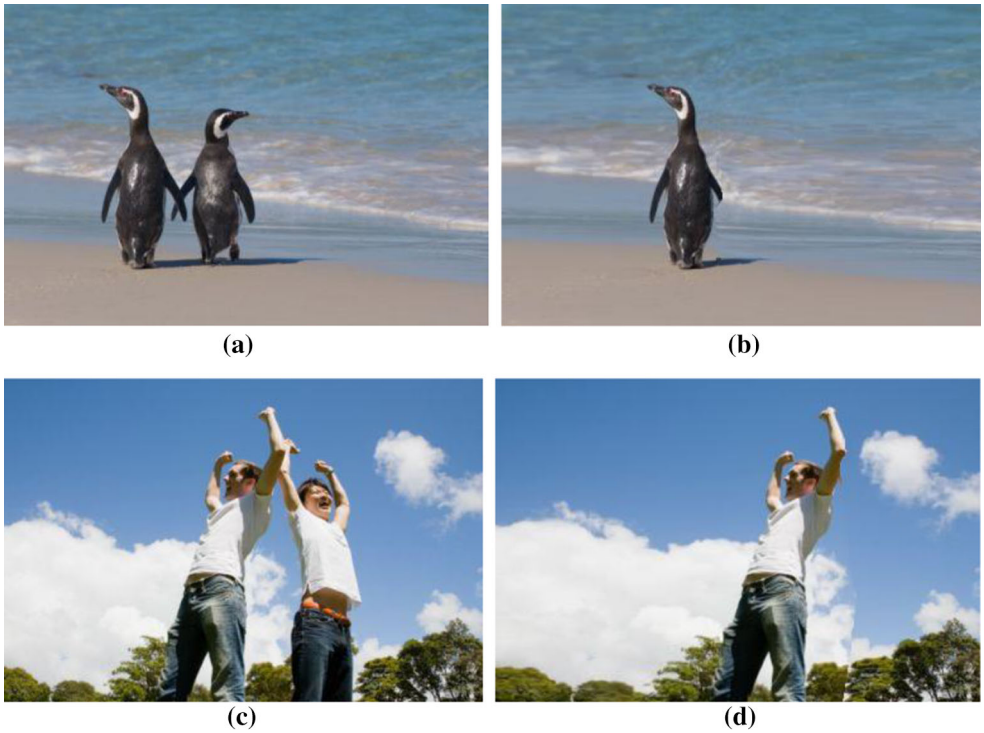

Fig. 2 Color images tampered by Seam-CI: **a**, **c** original image; **b**, **d** tampered image with some content removed

of the details, and image scaling stretched and distorted the details. By contrast, Seam-CI was able to retain most of the details. Figure 2 shows an overview of the proposed system, in particular it illustrates the system architecture.

Seam-CI can be used easily for image tampering, especially for the purpose of removing some unwanted details from a picture, as shown in Fig. 2.

The purpose of finding the "seam" is to enlarge or reduce an image without losing details when removing or inserting "seams". Therefore, the "seam" should meet two conditions: first, adding or deleting seams should not change the rectangular shape of the image; second, the pixels of seams should be as low energy as possible to ensure that image details are retained when adding or deleting seams.

To retain the rectangular shape of the original image, the seam needs to have only one pixel in every row (or column) of the image, and the adjacent pixels in the seam should be 8-connected. An $N_1 \times N_2$ image can be given as an example. Sorting all pixels of the original image by energy and then deleting the smallest N_1 pixels will most probably cause the amount of deleted pixels of every row to be different. As a result, the rectangular shape will not be retained. Similarly, sorting pixels of every row of the original image by energy and then deleting the smallest one from every row will most probably cause the image to be distorted by zigzag effects.

A vertical seam is defined as Eq. (3).

$$S^x = \{s_i^x\}_{i=1}^{N_1} = \{i, x(i)\}_{i=1}^{N_1},\ s.t. \forall i, |x(i) - x(i-1)| \leq 1 \tag{3}$$

where s represents the set of all pixels in a vertical seam. Figure 3 demonstrates the energy distribution in horizontal and vertical direction. Brighter ones means pixel possess more energy.

To minimize the energy sum of a seam, energy calculating methods for a pixel and a "seam" should be defined. The energy function of a pixel is

$$e_1(I) = \left|\frac{\partial}{\partial x} I\right| + \left|\frac{\partial}{\partial y} I\right| \tag{4}$$

where I is the intensity matrix of the original image. According to this function, the pixels in a "seam" can be represented as $I_s = \{I(s_i)\}_{i=1}^{N_1} = \{I(i, x(i))\}_{i=1}^{N_1}$. Therefore, the lowest-energy seam can be defined as

$$s^* = \min\_s\{E(s)\} = \min_s \left\{\sum_{i=1}^{N_1} e_1(s_i))\right\} \tag{5}$$

$$M(i, j) = e_1(i, j) + \min(M(i-1, j-1), M(i-1, j), M(i-1, j+1)) \tag{6}$$

From the second row to the last row, for a specific pixel (row i, col j) in a seam, three pixels can be connected: $(i-1, j-1)$, $(i-1, j)$, and $(i-1, j+1)$. A pixel is selected using Eq. (6).

where $M(i, j)$ is the energy sum of the pixels from first row to ith row; these pixels are all contained in the seam that passes (i, j). According to this rule, the energy sum can

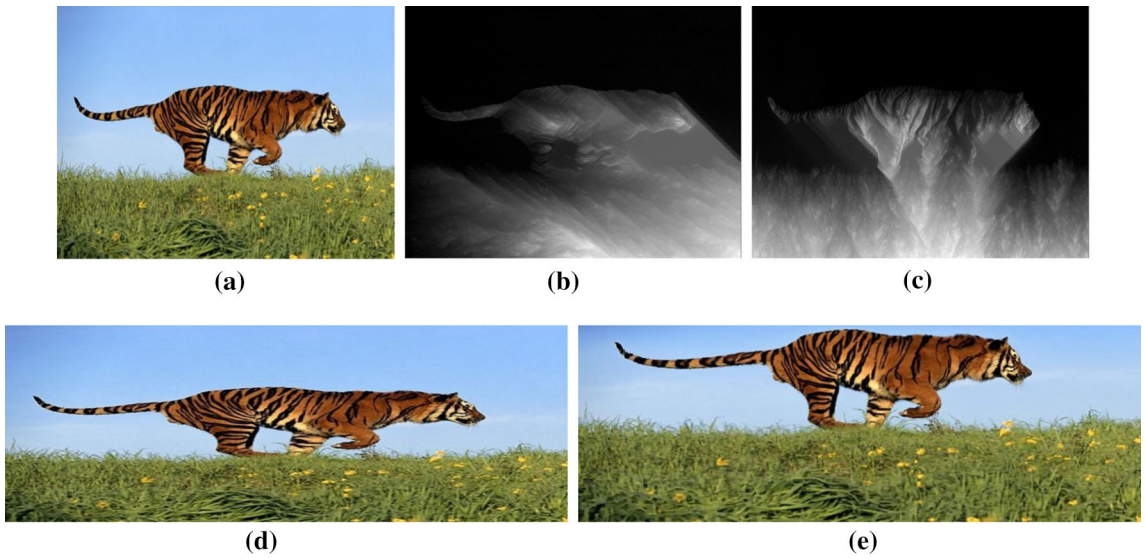

Fig. 3 Energy distribution and comparison of interpolation and Seam-CI: **a** original image **b** horizontal energy distribution **c** vertical energy distribution **d** enlarged image by interpolation both in horizontal and vertical direction **e** enlarged image by Seam-CI in same proportion as in **d**

be calculated row by row starting from the second row and ending in the last row. The $M(i, j)$ value of every pixel in the last row represents the energy sum of every alternative seam that can be sorted by energy to ensure that the low ones can be deleted or copied to resize an image. To find the exact position of a whole specific seam, backtracking should be used, starting from the pixels in the last row.

Image enlargement and reduction using Seam-CI are different. For image enlargement, seams should be inserted into the original image. The optimal vertical (horizontal) seam should be computed first. Then, the pixels of the seam are duplicated by averaging them with their left and right neighbors (top and bottom in the horizontal case). For a selected inserted seam, every pixel in the seam is replaced by two new pixels, which will be

$$b_1 = \text{round}\left(\frac{a_1 + a_2}{2}\right), \quad b_2 = \text{round}\left(\frac{a_2 + a_3}{2}\right) \tag{7}$$

where a_1, a_2, a_3 represents the original pixels in which a_2 is the pixel contained in the chosen seam, and a_1, b_1, b_2, a_3 represents the adjacent pixels after the seam is inserted. Clearly, a_2 has been replaced by b_1, b_2.

If the selected inserted pixel is located on the border, then it will be retained and only one new pixel will be inserted. For example, if a_1, a_2 are the adjacent pixels before seam insertion, then a_1 or a_2 belongs to the selected seam, and the adjacent pixels after inserting the seam will be a_1, b, a_2, where b is

$$b = \text{round}\left(\frac{a_1 + a_2}{2}\right) \tag{8}$$

3 Feature of DCT coefficients based on Benford's law

JPEG compression is block-DCT based and is the most popular image compression standard. For JPEG compression processing, the original image is first divided into non-overlapping 8×8 blocks. Two-dimensional DCT is then applied to each block. Then, the DCT coefficients are quantized using JPEG quantization table. Here, the term "DCT coefficients" is used to refer to the 8×8 DCT coefficients before quantization. Only the AC components of the DCT coefficients are considered in this work. The probability distributions of the DCT coefficients are usually modeled as Laplacian distribution or Cauchy distribution. A variety of applications based on these models have been proposed. However, the distributions of the most significant digits of the DCT coefficients have not been studied previously. The probability distribution of the first digits of the DCT coefficients will be studied in this section. Only 8-bit gray-level images are considered in this work for simplicity. The same principle can be extended to color images easily.

The widely used and recognized grayscale image set (Ng et al. 2004) from Columbia University is used to carry out the experiments. The Columbia image set contains 933 normal images and 912 splicing images; the 933 normal images are composed of 11 kinds of images that represent different texture distribution conditions. All images have the same size (128×128 pixels).

In Fig. 4, the mean distribution of the first digits of the DCT coefficients of 933 images in the Columbia database is shown by the yellow bars. The standard deviations of the distributions for these images are shown by the error bars (on

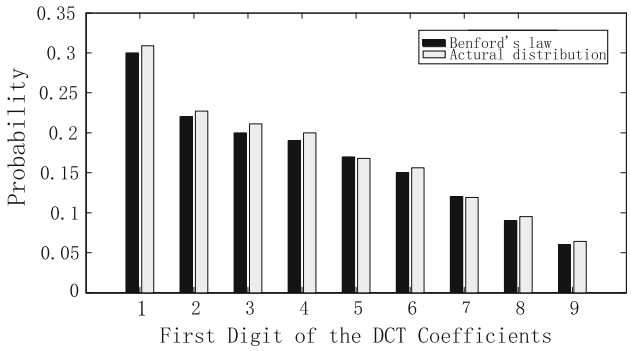

Fig. 4 Mean distribution of the first digits of the DCT coefficients of Columbia images

top of the yellow bars). Benford's law is also illustrated in red (left) bars for purposes of comparison. The probability distributions of the first digits of the DCT coefficients follow the standard Benford's law very well. The χ^2 divergence can be used to measure the quality of the fitting

$$\chi^2 = \sum_{i=1}^{9} \frac{(\hat{p}_i - p_i)^2}{p_i} \chi^2 \tag{9}$$

where $\hat{p}_i$ is the actual first digit probability, and p_i is the probability predicted by Benford's law as defined in Eq.(9), namely, $p_i = log_{10}(1 + \frac{1}{i})$. The average of χ^2 divergences for the fitting of all the Columbia images is only 0.0038, which indicates good fitting results.

SVM (Cortes and Vapnik 1995) can classify small samples and high-dimensional data efficiently. We use SVM to identify a Seam-CI tampered image from a normal image; SVM reduces the solution to a two-class classification problem. The LIBSVM (Chang and Lin 2011) that uses RBF kernel is used in our experiment.

4 Experimental process and result analysis

As previously mentioned, the Columbia image set, which contains 933 normal images and 912 splicing images, was used in this experiment. The 933 normal images contain the following 11 types of images that represent different texture distribution conditions:

- Texture (T): uniform texture images.
- Smoothing (S): smooth images.
- Texture-smoothing-vertical (TS-V): images have both texture parts and smoothing parts separated by a vertical boundary.
- Texture-smoothing-horizontal (TS-H): images have both texture parts and smoothing parts separated by a horizontal boundary.
- Texture-smoothing-others (TS-O): images have both texture parts and smoothing parts separated by an obvious boundary that is neither vertical nor horizontal.

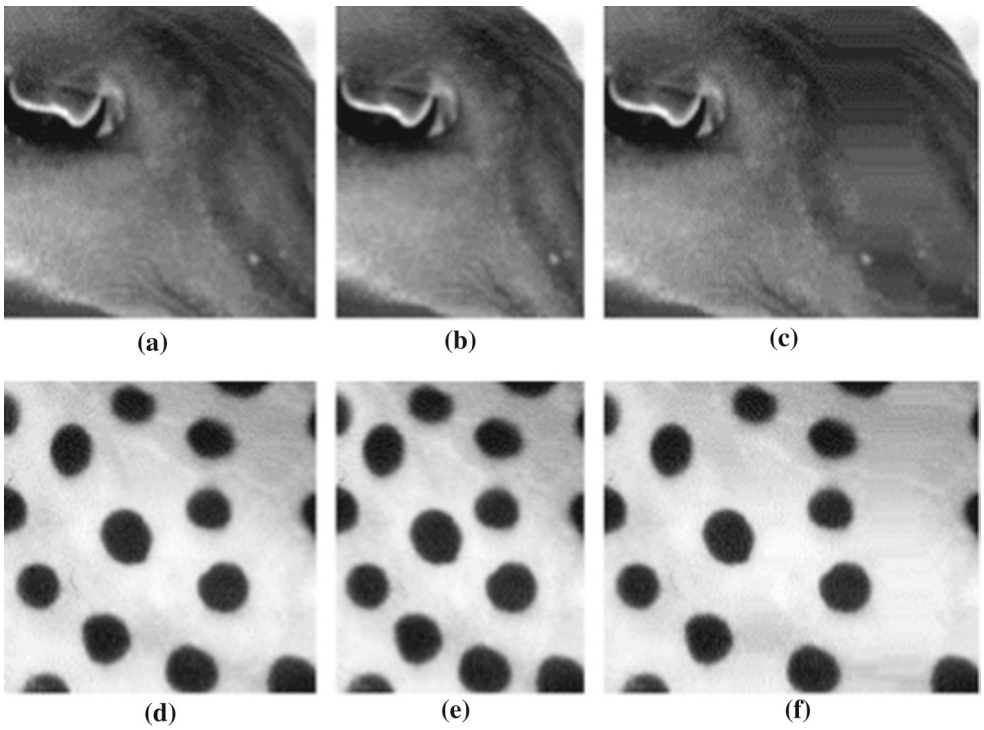

Fig. 5 Images tampered by Seam-CI **a**, **d** original image; **b**, **e** reduced image in 80 % proportion; **c**, **f** enlarged image in 120 % proportion

Table 1 Detecting result for gray-level image

	Au-S	Au-SS-H	Au-SS-O	Au-SS-V	Au-T	Au-TS-H
TPR (%)	99	98	97	94	95	96
TNR (%)	95	98	96	96	95	97
Accuracy (%)	97	98	96.5	95	95	96.5
	Au-TS-O	Au-TS-V	Au-TT-H	Au-TT-O	Au-TT-V	All
TPR (%)	94	98	99	92	92	95.8
TNR (%)	92	98	94	98	94	95.7
Accuracy (%)	93	98	96.5	95	93	95.75

- Texture-texture-vertical (TT-V): images have two texture parts separated by a vertical boundary.
- Texture-texture-horizontal (TT-H): images have two texture parts separated by a horizontal boundary.
- Texture-texture-others (TT-O): images have two texture parts separated by a boundary that is neither horizontal nor vertical.
- Smooth-smooth-vertical (SS-V): images have two smooth parts separated by a vertical boundary.
- Smooth-smooth-horizontal (SS-H): images have two smooth parts separated by a horizontal boundary.
- Smooth-smooth-others (SS-O): images have two smooth parts separated by an obvious boundary that is neither vertical nor horizontal.

In the experiments, we choose 933 normal images as the original set. The Seam-CI tampered image set is composed of enlarged images and reduced images. The enlarged images are obtained by performing seam insertion on 933 original images with 120 and 150% proportions. The reduced images are obtained by performing Seam-CI on the 933 original images with 80 and 50% proportion. Therefore, $4 \times 933 = 3732$ images are included in the tampered image set. The experiment on this set is credible because this set contains different texture distribution conditions. The images for training and testing are chosen from the 11 kinds of different texture distribution images. The training set is composed of 880 images; among the 440 normal images, 40 images are obtained from every kind of normal image; among the 440 tampered images, 40 images are obtained from every kind of tampered image, including 10 images each for 50, 80, 120, and 150% resizing ratios. In sum, the training set is composed of a total of 880 images, which include 440 normal images and 440 tampered images. Some sample images in the set are shown in Fig. 5.

Table 2 Comparison between algorithms by Sarkar and Fillion and the proposed method

	Sarkar	Fillion	Proposed method
TPR (%)	92	94	95.8
TNR (%)	91	92.7	95.7
Accuracy (%)	91.5	93.35	95.75

The testing set also contains 880 images; the selection strategy for these images is the same as introduced above

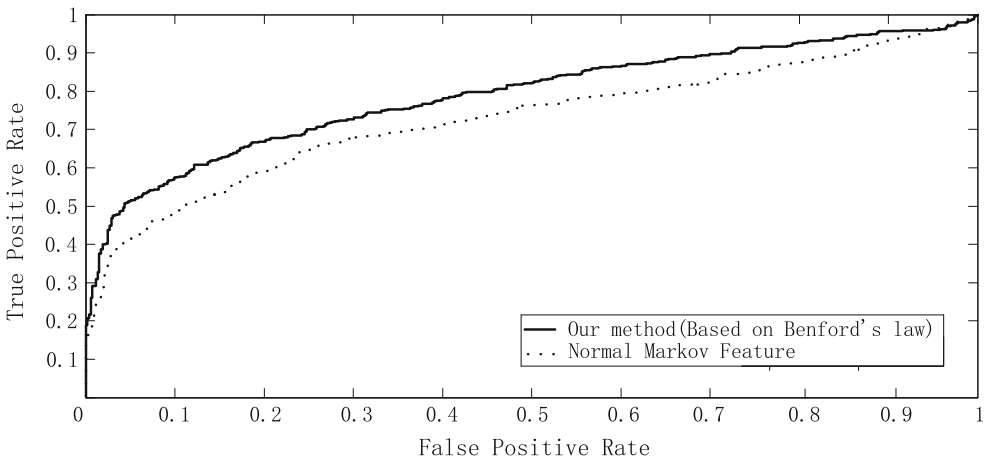

Fig. 6 ROC curve comparison between the Markov feature-based algorithm and the proposed algorithm

Table 3 Detecting result for color image

	CASIA
TPR (%)	99
TNR (%)	95
Accuracy (%)	97

except that all images in the testing set are different from the images in the training set.

The test result is shown in Table 1, where true positive rate (TPR) is the ratio between correctly detected tampered images and all tampered images to be detected; and true negative rate (TNR) is the ratio between the correctly detected normal images and all normal images to be detected. "Accuracy" represents the ratio between all the correctly detected normal and tampered images and all images to be detected. "All" represents the final detection result for all 11 types of images.

Table 1 shows that the detection accuracy, TPR, and TNR of the proposed method are all higher than 95 %.

We compared our algorithm with the fusion feature algorithm proposed by Fillion (2010). The comparison result is shown in Table 2. The experiments were conducted under the same environment. As shown in Table 2, the proposed method detects Seam-CI tampered images more precisely than the other methods.

The result of the comparison with the Markov feature-based algorithm can be revealed intuitively by the ROC curve shown in Fig. 6. The area under the curve of our algorithm is higher than that of the Markov-based algorithm, thereby confirming the superior performance of our algorithm.

To confirm, the proposed method also has excellent performance dealing with color images, we also perform experiments on color image dataset CASIA.

The CASIA 2.0 contains 7491 authentic and 5123 tampered color images. The images in this set are with different size, vary from 240 × 160 to 900 × 600 pixels. The set contains both uncompressed images and considered JPEG images with different compression ratio. The 7491 authentic images are all natural images and categorized into 9 categories: scene, animal, architecture, character, plant, article, nature indoor and texture. We choose 100 uncompressed images randomly from every category, get $9 \times 100 = 900$ authentic images all together, then perform Seam-Insertion on these images to get enlarged tampered images in 120 and 150 % proportions; perform Seam-Carving on these images to get reduced tampered images in 50 and 80 % proportions. Furthermore, we choose 1000 uncompressed images and tamper them by removing objects using Seam-Carving. Thus, $(900 \times 4) + 1000 = 4600$ tampered images are got to form tampered image set. Then all these 4600 tampered images be compressed into JPEG images with the ratio of 60, 80 % randomly. The training set and testing set are all from these 7491 normal images and 4600 compressed tampered images. Some sample images in the set are shown in Fig. 2. The detection result is shown in Tables 3 and 4.

5 Conclusion

A new algorithm based on Benford's law is proposed to detect content-aware image resizing using Seam-CI. The proposed algorithm established a statistic model based on Benford's law, which can be trained by SVM and used to classify a normal image from a Seam-CI-tampered image. Extensive experiments show that the detection accuracy of the proposed algorithm is superior to that of existing detection methods. The quantization step will distort the natural property of DCT coefficients, thereby resulting in a deviated distribution of first digit features. An expanded statistical model based on Benford's law may be studied in the future.

Acknowledgments This study was funded by National Natural Science Foundation of China (No. 61502218, 61232016, U1405254), Outstanding Young Scientists Foundation Grant of Shandong Province (No. BS2014DX016), Natural Science Foundation of Shandong Province (ZR2014FM005, ZR2013 FL008), Shandong Province Higher Educational Science and Technology Program (J14LN20), Shandong Province Science and Technology Plan Projects (2014 GGB01944, 2015GSF116001), Doctoral Foundation of Lu dong University (LY2014034, LY2013005, LY2015033), the Priority Academic Program Development of Jiangsu Higher Education Institutions (PAPD) and Jiangsu Collaborative Innovation Center on Atmospheric Environment and Equipment Technology (CICAEET), Guangzhou Scholars Project (No. 1201561613).

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Ethical approval This article does not contain any studies with human participants performed by any of the authors.

Table 4 Detecting result for nine categories of color images

	Scene	Animal	Architecture	Character	Plant	Article	Nature	Indoor	Texture
TPR (%)	98	97	98	94	94	95	98	95	93
TNR (%)	94	94	95	93	93	94	92	94	92
Accuracy (%)	96	95.5	96.5	93.5	93.5	94.5	95	94.5	92.5

References

Amerini I, Ballan L, Caldelli R, Del Bimbo A, Del Tongo L, Serra G (2013) Copy-move forgery detection and localization by means of robust clustering with j-linkage. Signal Process Image Commun 28(6):659–669

Amerini I, Ballan L, Caldelli R, Del Bimbo A, Serra G (2011) A sift-based forensic method for copy-move attack detection and transformation recovery. IEEE Trans Inf Forens Secur 6(3):1099–1110

Avidan S, Shamir A (2007) Seam carving for content-aware image resizing. In: ACM transactions on graphics (TOG), vol 26. ACM, New York, pp 10

Benford F (1938) The law of anomalous numbers. In: Proceedings of the American Philosophical Society, pp 551–572

Chang CC, Lin CJ (2011) Libsvm: a library for support vector machines. ACM Trans Intell Syst Technol (TIST) 2(3):27

Chi-Yao W, Hong Zhang Y, Chun Lin L, Wang SJ (2013) Visible watermarking images in high quality of data hiding. J Supercomput 66(2):1033–1048

Cortes C, Vapnik V (1995) Support-vector networks. Mach Learn 20(3):273–297

Fillion C, Sharma G (2010) Detecting content adaptive scaling of images for forensic applications. In: IS&T/SPIE electronic imaging. International Society for Optics and Photonics, San Francisco, pp 75410Z–75410Z

Fridrich AJ, Soukal BD, Lukáš AJ (2003) Detection of copy-move forgery in digital images. In: Proceedings of digital forensic research workshop, Citeseer

Fu D, Shi YQ, Su W (2006) Detection of image splicing based on Hilbert-Huang transform and moments of characteristic functions with wavelet decomposition. In: Digital Watermarking. Springer, Berlin, pp 177–187

Hill TP (1995) A statistical derivation of the significant-digit law. In: Statistical science, pp 354–363

Jin Li, Qian Wang, Cong Wang, Kui Ren (2011) Enhancing attribute-based encryption with attribute hierarchy. Mob Netw Appl 16(5):553–561

Li J, Chen X, Li M, Li J, Lee PC, Lou W (2014) Secure deduplication with efficient and reliable convergent key management. In: IEEE transactions on parallel and distributed systems **25**(6):1615–1625

Li J, Kim K (2010) Hidden attribute-based signatures without anonymity revocation. Inf Sci 180(9):1681–1689

Li J, Kim K, Zhang F, Chen X (2007) Aggregate proxy signature and verifiably encrypted proxy signature. In: Provable security. Springer, Berlin, pp 208–217

Li J, Wang Q, Wang C, Cao N, Ren K, Lou W (2010) Fuzzy keyword search over encrypted data in cloud computing. In: INFOCOM, 2010 Proceedings IEEE. IEEE, New York, pp 1–5

Lu W, Wu M (2011) Seam carving estimation using forensic hash. In: Proceedings of the thirteenth ACM multimedia workshop on multimedia and security. ACM, New York, pp 9–14

Sarkar A, Nataraj L, Manjunath BS (2009) Detection of seam carving and localization of seam insertions in digital images. In: Proceedings of the 11th ACM workshop on multimedia and security. ACM, New York, pp 107–116

Shi YQ, Chen C, Chen W (2007) A natural image model approach to splicing detection. In: Proceedings of the 9th workshop on multimedia & security. ACM, New York, pp 51–62

Simon Newcomb (1881) Note on the frequency of use of the different digits in natural numbers. Am J Math 4(1):39–40

Tian-Tsong N, Shih-Fu C, Sun Q (2004) A data set of authentic and spliced image blocks. Columbia University, ADVENT Technical Report, p 203

Tsung-Yuan Liu, Wen-Hsiang Tsai (2010) Generic lossless visible watermarking new approach. IEEE Trans Image Process 19(5):1224–1235

Wang J, Ma H, Tang Q, Li J, Zhu H, Ma S, Chen X (2013) Efficient verifiable fuzzy keyword search over encrypted data in cloud computing. Comput Sci Inf Syst 10(2):667–684

Yilei W, Wong DS, Zhao C, Xu Q (2015) Fair two-party computation with rational parties holding private types. Security and communication. Networks 8(2):284–297