Zhaofeng and Ming *EURASIP Journal on Image and Video Processing* (2018) 2018:69
https://doi.org/10.1186/s13640-018-0307-5

EURASIP Journal on Image and Video Processing

RESEARCH Open Access

Secure and efficient DRM watermark algorithm of forensics in mobile internet

Ma Zhaofeng[1*] and Jiang Ming[2]

Abstract

With the development of mobile Internet technology, the characteristic of easy editing, transmitting, and forging the digital media bring great challenges in authenticity of multimedia. Due to that, the focus on the digital forensics and identification, such as source identification, content authentication, and information integrity, have become an important research content in the field network security, forensic science and other fields. To solve this problem, in this paper, we proposed an image digital rights management scheme for mobile Internet forensics based on watermark with high-level security and traceability. In this scheme, we embed some user-identity-related and device-related information of mobile phones as a source identification of snapped photos and stored pictures; once the forwarded image data from mobile phones are misused such as spreading the image data on Internet without authorization, especially the confidential image data, we can trace the misuse responsibility by extracting the mobile phone data embedded in the mobile photos. Finally, we evaluate the proposed security and efficiency of the DRM watermark scheme for mobile Internet, a large amount of experiments manifest the proposed watermark scheme is robust, secure, and efficient for the protection and misuse tracing of mobile image data.

Keywords: Digital rights management, Watermark, Forensics

1 Introduction

Digital forensics is the technology of use of scientifically derived and proven methods toward the preservation, collection, validation, identification, interpretation, documentation, and presentation of digital evidence derived from digital sources for the purpose of facilitating or furthering there construction of events found to be criminal or helping to anticipate unauthorized actions shown to be disruptive to planned operations. The meaning of the term digital evidence refers to any information that can provide evidence, which contains stored or digitally transmitted data, and they can be used in court as evidence.

In the past few years, much more digital evidences have been used in cases, such as documents, e-mail, digital photos, message history, historical records database of browser, computer memory data, GPS tracking data, and digital images or sound files.

With the rapid development of multimedia, especially the popularity of the mobile and the convenience of the mobile Internet, much more image data produced from mobiles are much easier to spread in the circulation of the mobile Internet. According to statistics, by the end of June 2017, the total number of mobile phone users in China is 1 billion 360 million, and the Internet users of mobile phones break through 1 billion 100 million. The configuration of mobile camera is more advanced, and people use the portable mobile phone much more conveniently to get high-quality photos of a scene in life. The close connection between mobile phones and people's lives makes much more cases based on the evidence of mobile data happened in the judicial practice. Digital image forensic techniques are the technologies and methodologies for judging the origin, integrity, and authenticity of digital images, by analyzing, identifying, and authenticating the steganography, forgery, and tampering of digital images. There are two main research directions in digital image forensic techniques [1]: image source detection and tamper detection [2].

For mobile image source detection, the lens and imaging sensors of different brands and models of mobiles

* Correspondence: mzf@bupt.edu.cn
[1]Information Security Center of Beijing University of Posts and Telecommunications, Beijing 100876, China
Full list of author information is available at the end of the article

are different, with which the process of digital images are also different, including compression and storage. Therefore, we extract and analyze the small differences, such as image style and quality, to identify the source of the digital mobile images. There are three main types of detection technologies for image sources forensics: (1) Mobile digital image sources are taken as the evidence based on mobile device type: we use a classifier to classify mobile image sources by extracting statistical characteristics, such as image quality, color, lens radial distortion, and wavelet coefficients. Levent Özparlak proposed statistical image models for wavelet-based transforms and compared their relative merits within the context of digital image forensics [3]. Farid described a series of psychophysical experiments that used images of varying resolutions, JPEG compression, and colors to explore the ability of observers to distinguish computer-generated from photographic images of people [4]. The researchers also put forward some effective features to distinguish and identify the digital mobile and CG images, such as CFA interpolation periodicity [5], visual characteristics [6], texture feature [7, 8], white balance [9], and local binary pattern(LBP) [10]. For more different types of mobile devices, Orozco used color features and quality features [11]. (2) Mobile digital image sources are taken as the evidence by mobile device model: the same tomography hardware and the same image processing algorithm is used in mobiles with the same device model, so we use the characteristics of hardware device and image processing algorithm to be the important foundation for mobile image forensics. Chen [12] estimated the interpolation coefficient of CFA and selected a suitable classifier to achieve higher accuracy of source identification. White balance [13] is an important image post-processing algorithm in mobile and camera imaging system, and its parameter estimation is also applied to identify the source of digital image. Another part of research work takes the entire mobile image acquisition equipment as a whole and expects to build a whole model to describe the difference in the mobile image equipment from different angles and realize the source forensics. Goyal [14] evaluates the effectiveness of image quality measures (IQM) for identifying the source mobile phone from the images or videos captured by that mobile phone. A new single-image feature for mobile identification is proposed based on a local binary pattern of extracted edges form the input mobile image [15]. Thai proposed a statistical approach for the camera model identification problem, which is based on the heteroscedastic noise model for describing a natural raw image [16]. Unlike the preceding methods, Luan takes the source forensics as a clustering problem in unsupervised learning. To avoid using any prior knowledge in practical scenarios, Luan proposed a graph-based approach to classify the source cell-phones [17]. (3) Mobile digital image sources are taken as the evidence by mobile device individual characteristics: we can determine the mobile image source by correlation detection between the pattern noise and abnormal pixels caused by inherent defects of the imaging device. Fridrich [18] divides sensor mode noise into inherent mode noise and optical response non-uniformity noise, obtains the mode noise of digital camera by filtering and statistical difference, and realizes the forensics of camera individual sources of digital images using correlation detection, hypothesis testing, and other methods. But besides that all, sensor mode noise are also used for device individual source forensics in mobile [19], portable digital video camera [20], and scanner [21].

For mobile image tamper detection, the tamper and the collector of the mobile images do against with each other. The tamper is to make false images of truth as fast as you can, while the collector is to try the best to find evidence of the image being tampered with. Although it is difficult to be easily discovered, the tampered image causes more or less damage to the inherent continuity of the natural image during the capturing process and storing process. Therefore, the collectors use some statistical characteristics of natural images to detect the types of image tampered and display the detailed location of tampering. There are also three main types of detection technologies for image tamper forensics. (1) The first method is based on resampling, such as copy-paste [22, 23] and splice-synthesis [24]. (2) The second method is based on compression, such as dual compression [25] and JPEG blocking artifact [26, 27]. (3) The third method is based on imaging device features, such as image statistical characteristics [28], pattern noise [29] and CFA interpolation [30–35].

2 Methods

2.1 Watermark algorithm principle for source forensic and tamper detection

In order to solve the problem of source detection and tamper detection of mobile image data, we proposed a new watermark-based mobile image data security scheme to protect confidential image data in this paper. Robust watermark, such as mobile-identification-related and user-identification-related information, are used as in this scheme for source detection, while the features of image itself are used as semi-fragile watermark in this scheme for traceability. Finally, for security and efficiency, we evaluated the proposed mobile image data security scheme through several sets of experiments.

As the storage capacity of the mobile phone is limited, the image data is stored and transmitted in a compressed way. Considering JPEG picture compression format is widely used in mobile phones, we use JPEG as an example of confidential image data in our proposed mobile image data security solution. In this scheme, robust watermark for source forensics is a binary image, which indicates some text info related to mobile identification and user identification. Semi-fragile watermark data for tamper detecting is also a binary image, which is generated by the host image itself. In order to avoid the impact between the two digital watermarks, the two watermarking algorithms should be independent. Considering JPEG compression is lossy, which may bring some errors to watermarks, the watermark scheme should be robust to JPEG compression at least. Discrete cosine transform (DCT) is reversible in JPEG, and the DCT coefficients have a certain degree of independence, so we choose to embed watermarks based on the DCT coefficients in this scheme.

2.2 JPEG image compression and DCT transform

JPEG compression usually divides the completely digital image into 8 × 8 blocks, based on DCT, and quantifies the luma samples of each block (see Fig. 1). The DCT of two-dimensional signal $f(i, j)$, $i = 0, 1, \ldots M-1$; $j = 0, 1, \ldots N-1$) f (i, j), i = 0, 1, …M − 1; j = 0, 1, …N − 1) is defined as follows:

$$F(u,v) = c(u)c(v)\sum_{i=0}^{M-1}\sum_{j=0}^{N-1} f(i,j)\ cos\frac{(2i+1)u\pi}{2M}\ cos\frac{(2j+1)v\pi}{2N}$$

$$u = 0, 1, \ldots, \mathrm{M}{-}1; v = 0, 1, \ldots, \mathrm{N}{-}1. \quad (1)$$

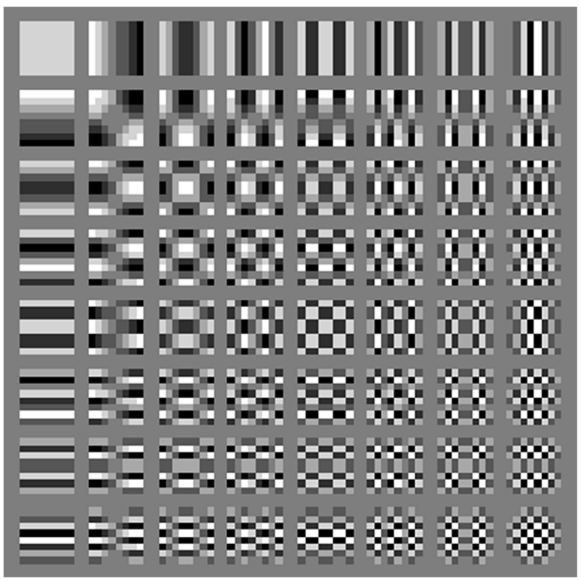


Fig. 1 DCT transform

The inverse IDCT is defined as follows:

$$f(i,j) = \sum_{u=0}^{M-1}\sum_{v=0}^{N-1} c(u)c(v)F(u,v)\ cos\frac{(2i+1)u\pi}{2M}\ cos\frac{(2j+1)v\pi}{2N}$$

$$u = 0, 1, \ldots, \mathrm{M}{-}1; v = 0, 1, \ldots, \mathrm{N}{-}1. \quad (2)$$

$$c(u) = \begin{Bmatrix} \frac{1}{\sqrt{M}} & u = 0 \\ \frac{2}{\sqrt{M}} & u \neq 0 \end{Bmatrix} c(v) = \begin{Bmatrix} \frac{1}{\sqrt{N}} & v = 0 \\ \frac{2}{\sqrt{N}} & v \neq 0 \end{Bmatrix}$$

2.3 JPEG quantization

JPEG encodes DC coefficient using differential pulse code modulation and encodes AC coefficients in a zigzag sequence, as shown in Fig. 2a. Figure 2b is the quantization table for JPEG compression.

2.4 Masking model of HVS

The main idea of the adaptive watermarking method based on human visual mode is HVS that can help us embed watermarks in different regions with different intensities adaptively, which means we focus on minimizing watermark strength in regions of image sensitive to HVS and maximizing watermark strength in regions of image not sensitive to HVS. Texture masking and illuminance masking indicate we should embed watermark intensively in the regions of images with high texture and brightness.

Performing two-level wavelet transform on the image, the quad tree structure is shown in Fig. 3. The model of the illuminance sensitivity is shown in Fig. 4.

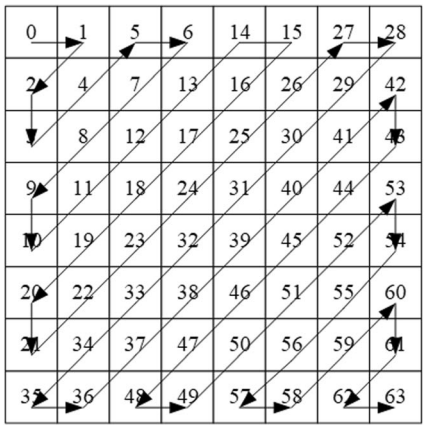
Supposing two-dimensional image is $f(x, y) \in L^2$ (R^2), two-dimensional wavelet decomposition image is:

$$A_j f = A_{j+1} f + D^2_{j+1} f + D^3_{j+1} \quad (3)$$

$$A_{j+1} f \sum_{m=-\infty}^{\infty}\sum_{n=-\infty}^{\infty} C_{j+1}(m,n)\varnothing_{j+1}(m,n)D^i_{j+1} = \sum_{m=-\infty}^{\infty}\sum_{n=-\infty}^{\infty} D^i_{j+1}(m,n)\varnothing^i_{j+1}(m,n)\ i = 1, 2, 3 \quad (4)$$

$$\begin{cases} C_{j+1}(m,n) = \sum_{k=-\infty}^{\infty}\sum_{l=-\infty}^{\infty} h(k-2m)h(l-2n)C_j(k,l) \\ D^1_{j+1}(m,n) = \sum_{k=-\infty}^{\infty}\sum_{l=-\infty}^{\infty} h(k-2m)h(l-2n)C_j(k,l) \\ D^2_{j+1}(m,n) = \sum_{k=-\infty}^{\infty}\sum_{l=-\infty}^{\infty} h(k-2m)g(l-2n)C_j(k,l) \\ D^3_{j+1}(m,n) = \sum_{k=-\infty}^{\infty}\sum_{l=-\infty}^{\infty} h(k-2m)g(l-2n)C_j(k,l) \end{cases} \quad (5)$$

where $h(\cdot)$ and $g(\cdot)$ are the horizontal filter function and



0	1	5	6	14	15	27	28
2	4	7	13	16	26	29	42
3	8	12	17	25	30	41	43
9	11	18	24	31	40	44	53
10	19	23	32	39	45	52	54
20	22	33	38	46	51	55	60
21	34	37	47	50	56	59	61
35	36	48	49	57	58	62	63

8	6	5	8	12	20	26	30
6	6	7	10	13	29	30	28
7	7	8	12	20	29	35	28
7	9	11	15	26	44	40	31
9	11	19	28	34	55	52	39
12	18	28	32	41	52	57	46
25	32	39	44	52	61	60	51
36	46	48	49	56	50	52	50

Fig. 2 **a** and **b** is the DTC quantization for JPEG compression

vertical filter function, respectively. j means the composition level.

Texture masking matrix is $M \times N$, shown as Ts .

$$Ts(m,n) = \delta(\mathrm{LH1}) + \delta(\mathrm{HL1}) + \delta(\mathrm{HH1}) \tag{6}$$

$$\delta = \sqrt{\frac{1}{M \times N}\sum_{m=0}^{M-1}\sum_{n=0}^{N-1}(x(m,n)-\mu)^2} \tag{7}$$

$$\mu = \frac{1}{M \times N}\sum_{m=0}^{M-1}\sum_{n=0}^{N-1}x(m,n)^2 \tag{8}$$

$$m = 0,1,2,\ldots,\mathrm{M}-1; n = 0,1,2,\ldots,\mathrm{N}-1.$$

The normalization of Ts is recorded as TM .

Normalizing the LL2 coefficient matrix, illuminance masking matrix is expressed as Ls. The size of matrix LL2 is $M \times N$,

$$L(m,n) = f(\mathrm{LL2}(m,n)) \tag{9}$$

$$f(x) = \begin{cases} -\frac{1}{180}x + 3.5 & 0 \le x < 90 \\ -\frac{3}{38}x + \frac{192}{19} & 90 \le x < 128 \\ \frac{2}{127}x - \frac{256}{127} & 128 \le x < 255 \end{cases} \tag{10}$$

$$m = 0,1,2,\ldots,\mathrm{M}-1; n = 0,1,2,\ldots,\mathrm{N}.$$

where $f(\cdot)$ is a function of the model of Illuminance sensitivity, and normalization of Ls is recorded as LM .

The masking matrix is as follows:

$$SS = a \cdot LM + b \cdot TM \tag{11}$$

where $a = 1 - b$. Normalization of SS is recorded as SM.

Figure 5 reflects the much higher brightness part of the image, the much more suitable for embedding watermark.

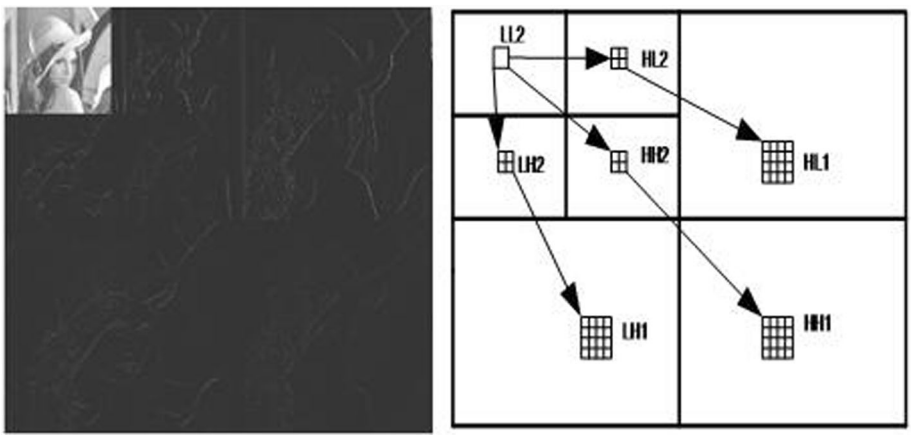


Fig. 3 The quadtree structure of DWT for Lena is shown in Fig. 3

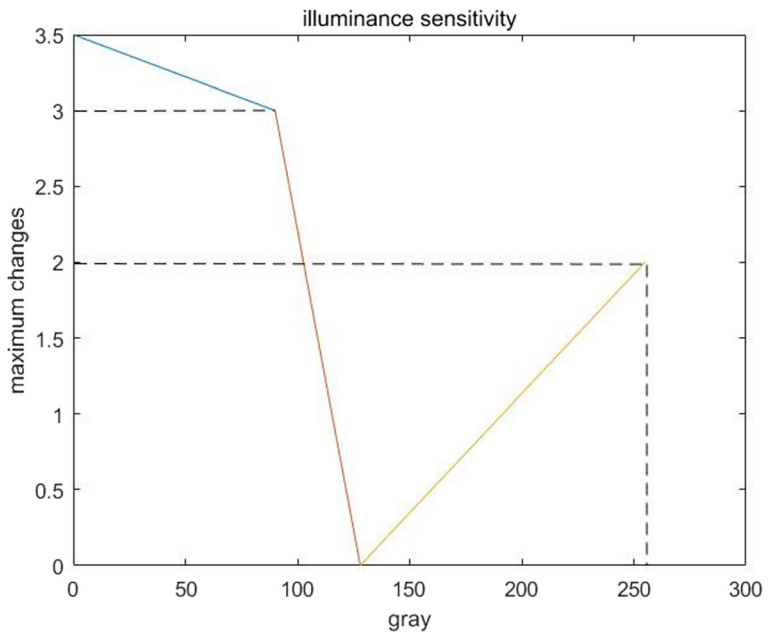


Fig. 4 Model of illuminance sensitivity

JPEG standard is based on 8×8 blocks, and the masking of visual for block image is calculated out by the average masking of pixels of block image, recorded as Q.

2.5 The proposed mobile image data security scheme based on watermark

2.5.1 Watermark scheme architecture for mobile internet forensics

Watermark scheme architecture for mobile internet forensics is shown in the following Fig. 6, which includes the source forensic watermark algorithm, tamper tracing watermark algorithm, and tamper detecting algorithm.

2.5.2 Watermark embedding algorithm for source forensics

In the watermark embedding process, we choose one $M \times N$ binary image as watermark W. $W = \{W(i,j) | \, 0 \le i < M, 0 \le j < N\}$, and $W(i,j) \in \{0, 1\}$. We scramble the binary image for security and then convert the image into a one-dimensional watermark sequence, namely $W = \{w_i\}$, $i = 1, 2, \ldots, C$; $C = M \times N$, $w_i = 0$ or 1. The host image is divided into 8×8 blocks, named X_i, and $x_i(m, n)$ is the pixel value in (m, n) of X_i. The sequence of DCT coefficients of one block is recorded as $C_i(j)$, $(j = 0, 1, 2, \ldots 63)$ in a zigzag scan form. i is the number of blocks. We choose a continuum of mid-frequency coefficients to embed the watermark, such as

$$C_i(k-2), C_i(k-1), C_i(k), C_i(k+1), C_i(k+2)$$

$$k = 2, 3, \ldots, 61$$

The processing of coefficient $C_i(k)$ balances the transparency and robustness of watermark. The specific methods are as follows:

if $w_i = 0$, then we can get

Fig. 5 Texture image of Lena. $a = 0.1$. $a = 0.3$. $a = 0.5$. $a = 0.7$. $a = 0.9$

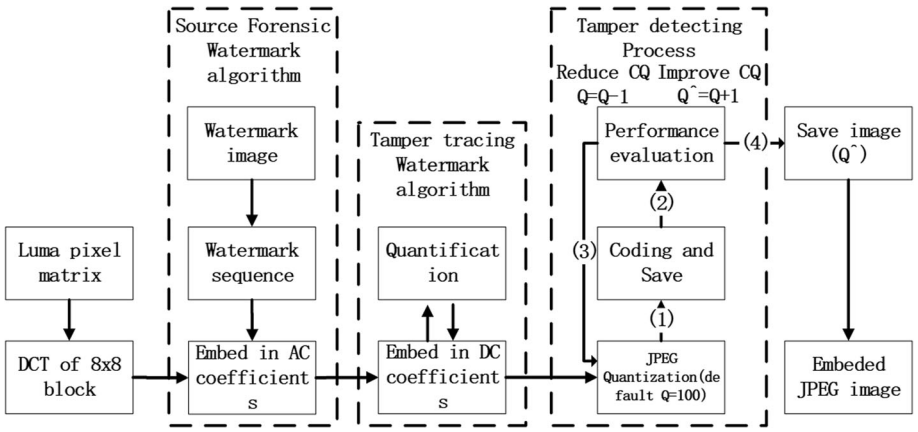


Fig. 6 Flow of watermark embed scheme

$$C_i(k)' = MIN(C_i(k-2), C_i(k-1), C_i(k), C_i(k+1), C_i(k+2)) - Q_i \quad (12)$$

else if $w_i = 1$, then we can get

$$C_i(k)' = MAX(C_i(k-2), C_i(k-1), C_i(k), C_i(k+1), C_i(k+2)) + Q_i \quad (13)$$

$C_i(k)'$ is the modified coefficient, MIN(.) denotes the minimum value function, MAX(.) denotes the maximum function. and Q_i can control the watermark strength. Q_i is defined in Fig. 7.

Five consecutive DCT coefficients are being as host sequences, and how to select the coefficients is the key to the algorithm, which directly affects the performance of the watermark. As shown in Fig. 1, different $C_i(k)$ means different impacts on block images. In order to balance the transparency and robustness of the watermark in this algorithm, we select medium- and low-frequency coefficient watermark to embed. Experimental analysis of different $C_i(k)$ and PSNR has been carried out in Fig. 8. The host image is the standard image Lena, and the watermark is from a pseudo-random sequence.

In the graph, the horizontal ordinate is the location of the watermark coefficient that is the k of $C_i(k)$. The vertical ordinate is the PSNR of watermarked JPEG image. Overall, with the offset of the embedded point, embedding a watermark in low frequency may bring low imperceptibility, while embedding watermark in high frequency may bring better visual quality of the watermarked image. The maximum value appears in the curve line when k is 12, in which the PSNR is 42.78 dB. The main reasons may be as follows:

1. From Fig. 1, we can see the change of medium frequency $C_i(12)$ impact the block image evenly.

Fig. 7 The masking characteristic of Image

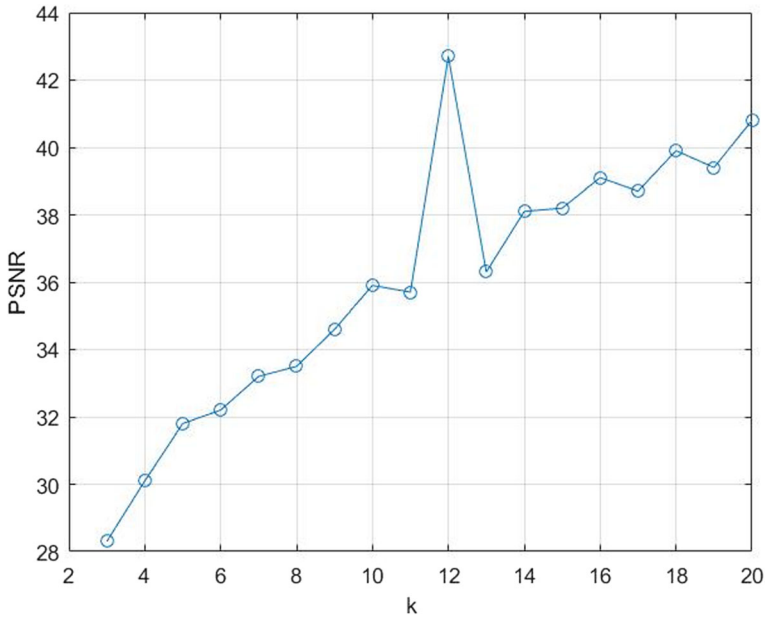


Fig. 8 PSNR for different *k*

2. When *k* is 12, the modification of the coefficient is in relation to five consecutive coefficients, such as $C_i(10)$, $C_i(11)$, $C_i(12)$, $C_i(13)$, and $C_i(14)$, which are in the same layer in the zigzag scanning line and very even.
3. As show in Fig. 2b, the quantization step is 8 when the *k* is 12, which is smaller than the steps around, so the quantization errors may be smaller and bring high PSNR of block images.

From the above analysis, we select $C_i(12)$ $(k = 12)$ as the watermark embedding point.

From the formula 11, we can get Q_i that is related to the texture of the image. The much bigger Q_i means the higher texture in the blocks of the image and the easier to embed watermarks. The texture image of Lena using different *b* is as shown in Fig. 5.

2.5.3 Watermark embedding algorithm for tamper tracing

Except for source forensics, the scheme also needs to embed another watermark for tamper tracing of the mobile images on the Internet. Because the host image will be saved in JPEG, the semi-fragile watermark for tamper tracing should be robust to JPEG compression. To avoid attacking the watermark for source forensics embedded in DCT AC mid-frequency coefficients, the watermarks for tamper tracing are chosen to be embedded in DCT DC coefficients. For blind watermarking algorithm, a method of coefficient quantization will be used. The procedures are as follows:

Step 1**:** The algorithm creates a two-valued pseudo-random sequence using the length and width of the host image.

$$S = \text{Random}\ (M, N, k) \tag{14}$$

M and N represent the length and width of the image, respectively. k is a key for random function. S is a two-value sequence with only 0 and 1 (s_i = 0 or 1), the length of which is the number of block images. Random (.) is the function for pseudo-random.

Step 2**:** And then,

$$C_i(0)' = \begin{cases} \text{if } s_i = 0 \\ Q(C_i(0)) \times \gamma + \left(\left\lceil \frac{\gamma}{2} \right\rceil + mod(C_i(0), 2) \times \gamma\right) + 1.5 \times \gamma \\ \text{else if } s_i = 1 \\ Q(C_i(0)) \times \gamma + \left(\left\lceil \frac{\gamma}{2} \right\rceil + mod(C_i(0) + 1, 2) \times \gamma\right) + 1.5 \times \gamma \end{cases}$$

$$Q(C_i(0)) = \left\lfloor \frac{C_i(0) - 1}{\gamma} \right\rfloor \tag{15}$$

In the formula above, $C_i(0)'$ is the modified coefficient after watermarking algorithm, $C_i(0)$ is the DC coefficient of every 8×8 block image, γ describes the quantization steps, *i*is the sequence number corresponding to the *i*th block of the image. In order to detect the tamper region in the entire image, all blocks of images should be embed the watermark using the formula 15.

2.5.4 Tamper detecting process of watermark scheme

After the source forensic algorithm and tamper tracing algorithm, the DCT coefficients are compressed into JPEG images after quantization and coding. In the scheme, the forensic algorithm is robust and the tamper tracing algorithm is semi-fragile. Lossy JPEG compression may cause some errors in watermark extracting, which should be verified that has no effects on the tamper tracing detecting. The low compression rate causes the size of the watermarked JPEG images to be too large especially for mobile storage; meanwhile, the high compression rate may cause some errors in tamper detection. Therefore, in the scheme, we carry out one tamper detecting process to choose the most suitable quality factor for JPEG compression, which is very important for the size of JPEG saving images. The tamper detecting process is described as follows:

Step 1: Store mobile JPEG image watermarked with quality factor $Q = 100$. The PSNR of the watermarked JPEG image is very high, but the size of image is too large. The choice of the quality factor is very important for the watermark performance, so the performance evaluation will be go on next;

Step 2: Conduct performance evaluation on watermarked JPEG image. Performance evaluation mainly considers three aspects: tampering, the CIR (formula 20), and the bit error rate of the robust watermark for source forensic. If there is no tamper, the CIR and bit error are in a certain threshold range, go to step 3, or else go to step 4; the specific tamper detect methods are as follows:

Read watermarked JPEG image, based on DCT, and quantify the luma samples of each block:

1. : We can get s_i like the method as formula 14.
2. : And then,

$$s_i' = \begin{cases} 0 & mod(Q(C_i(0)), 2) = 0 \\ 1 & otherwise \end{cases}$$

$$Q(C_i(0)) = \left\lfloor \frac{C_i(0) - 1}{\gamma} \right\rfloor \tag{16}$$

if $(s_i' == s_i)$, there is no tamper in block image, else the block is tampered.

$C_i(0)$ is the DC coefficient in the ith 8×8 block image, $F(.)$ indicates the rounding down function, and γ describes half the quantization step. If the tamper block is needed to be labeled, we can write the block image black.

Step 3: Since there is no tampering, we can save the watermarked JPEG image with the new compression quality $Q = Q - 1$. And then, the size of the image may be smaller, which is very important for file storing in a mobile. After that, we also should verify the tamper detection, go to step 2;

Step 4: Once any block image is tampered, which shows that the quality factor of JPEG compression is too small and the lossy compression affects the tamper detection of the watermark, the quality factor needs to be improved. Store the watermarked JPEG image with the quality factor $Q = Q + 1$, then the whole watermark embed scheme is complete.

2.5.5 Watermark extracting algorithm

2.5.5.1 Watermark extracting algorithm for source forensic In order to take the evidence for images from the mobile Internet, we should extract the forensic watermark in the suspicious target image. After getting the JPEG images from a mobile, we still need to divide the JPEG image into 8×8 blocks, based on DCT, and quantify the luma samples of each block, and we also get the DCT coefficients $C_i(j)$, $(j = 0, 1, 2, \ldots 63)$ the same as the watermark embedding process, i is the sequence number corresponding to the ith block of the image. We select the mid-frequency coefficients to extract watermarks, such as $C_i(k-2)$, $C_i(k-1)$, $C_i(k)$, $C_i(k+1)$, $C_i(k+2)$ corresponding with the embed process.

Extracting method is as follows:

$$\text{if } \mathrm{C_i(k)} > \Big(\mathrm{C_i(k{-}2) + C_i(k{-}1), + C_i(k) + C_i(k+1)} + \mathrm{C_i(k+2)}\Big)/5, \mathrm{w_i} = 1.$$

else

$$w_i = 0.(17)$$

where w_i is the ith watermark bit. At last, anti-scramble the watermark bit to get the watermark image.

2.5.5.2 Watermark extracting algorithm for tamper detecting The process of tamper detection is described as step 2 in Section 2.5.4. If one block image is tampered, mark it with black.

3 Results and discussions

3.1 Experiments and evaluations of the DRM watermark algorithm

In this scheme, we designed watermark-based effective digital rights management (DRM) algorithm for mobile

Internet forensics by embedding watermarks based on DCT, in which the robust watermark is for forensics while semi-fragile watermark is for tamper tracing. In our scheme, we constructed a watermark in a figure format rather than in a text format; thus, even if the image data is attacked much more, the watermark can still be recognized the figured watermark and can parse the user-related and hardware-related information and then can take the evidence by the figured watermark.

In order to evaluate the scheme, we carry out various experiments. In terms of watermark invisibility, PSNR is used for evaluating image quality objectively, which is defined as:

$$PSNR = 10\ \log_{10}\left(\frac{(O-1)^2}{MSE}\right)dB \tag{18}$$

where O – 1 represents the maximum value of the original image pixels. MSE is the mean squared errors, given by

Table 1 Different quality factors for stadium image

Quality factor	CIR (%)	PSNR (dB)	Bit error (%)	Is tampered?
100	169.69	39.00	0	NO
95	42.25	39.03	0	NO
90	11.26	38.64	0	NO
85	− 10.7	38.99	0.0001	NO
80	− 18.34	38.44	0.0004	NO
75	− 23.57	37.80	0.0010	NO
70	− 28.26	37.36	0.0017	NO

$$MSE = \frac{1}{M \times N}\sum_{i=0}^{M-1}\sum_{j=0}^{N-1}\left[W(i,j) - \hat{W}(i,j)\right]^2 \tag{19}$$

In this experiment, some photos taken by a mobile phone with the size of 3120 × 4160 is used. A binary logo of size 250 × 250 is used as a watermark, which describes some info for the mobile phone, such as phone no., time, location and IMEI. Figure 9 shows the original image, binary watermark, and the corresponding watermarked

Fig. 9 Watermark embedding and extracting

image. The PSNR of the watermarked image is 38.87 dB while the quality factor is 93.

After the watermark is embedded, the watermarked JPEG file storage capacity may be increased. To control storage capacity, the watermarked JPEG image should be compressed and stored at a low quality factor. We use the index CIR (capacity increase ratio) as the percentage of image capacity increase-ratio between the host JPEG image and the watermarked JPEG image.

Fig. 10 Watermark attacking and detecting

(a) Stadium image

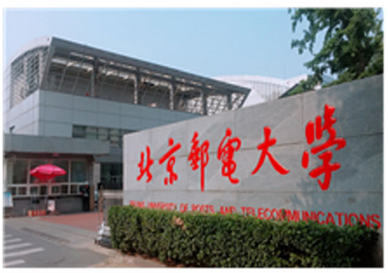
(b) Binary watermark

(c) Motto image

(d) Binary watermark

(e) Research Building image

(f) Binary watermark



(g) Science park image

(h) Binary watermark

Fig. 11 Different mobile pictures with watermarks

$$\mathrm{CIR} = \frac{\mathrm{Capacity(watermarked)} - \mathrm{Capacity(original)}}{\mathrm{Capacity(original)}} \times 100\% \tag{20}$$

where Capacity(watermarked) and Capacity(original) are the file capacity of the watermarked image and the original image capacity.

We can analyze the following conclusions from Table 1.

The CIR of JPEG image decreases when the quality factor decreases. Even when the quality factor is below 90, CIR is lower than 0. When compression factor is reduced to 90, bit error happens. In tamper detecting, even when the quality factor goes down to 70, no tamper happens.

3.2 Attacks and analysis of DRM watermark algorithm

In this experiment (see Figs. 10, 11 and 12), we firstly attack the watermarked image and then verify whether we can extract the watermark from the object image attacked and also whether we can track the tampered tracks. In the experiment, we gave the detailed experiments according to the four kinds of attacks: (1) copy and paste attack, (2) insert a picture attack, (3) insert graph attack, and (4) insert color inverse attack, where the NC of the extracted watermark is 0.9787, 0.9945, 0.9891, and 0.9554, respectively.

For further experimental verification, we use different mobile images as experimental objects. The watermark describes the mobile photo information, such as phone no., time, location, and IMEI, and the attack methods are the same as above. Table 2 shows the same results as described above and the experiments show the validation and performance of the proposed algorithm.

From the results in Fig. 9, the watermark is easily recognized, which describes the mobile photo information, such as phone no., time, location, and IMEI after the attacks.

According to the experimental results from Table 2, we draw the conclusion as follows: The CIR for different images are in general less than 0 when compression quality is 90, which means the file size of watermarked images cannot be increased very much for watermark embedding, even decreased for JPEG compression. The PSNR of the images are in general above 38 dB, and the bit error of the identification watermarks are near 0.

The related methods are compared with the proposed schemes in terms of functionalities and performance.

Comparing with current related schemes, our proposed scheme has the following three advantages and innovations.

Obviously, in addition to the above-mentioned analysis, the present schemes have another three advantages.

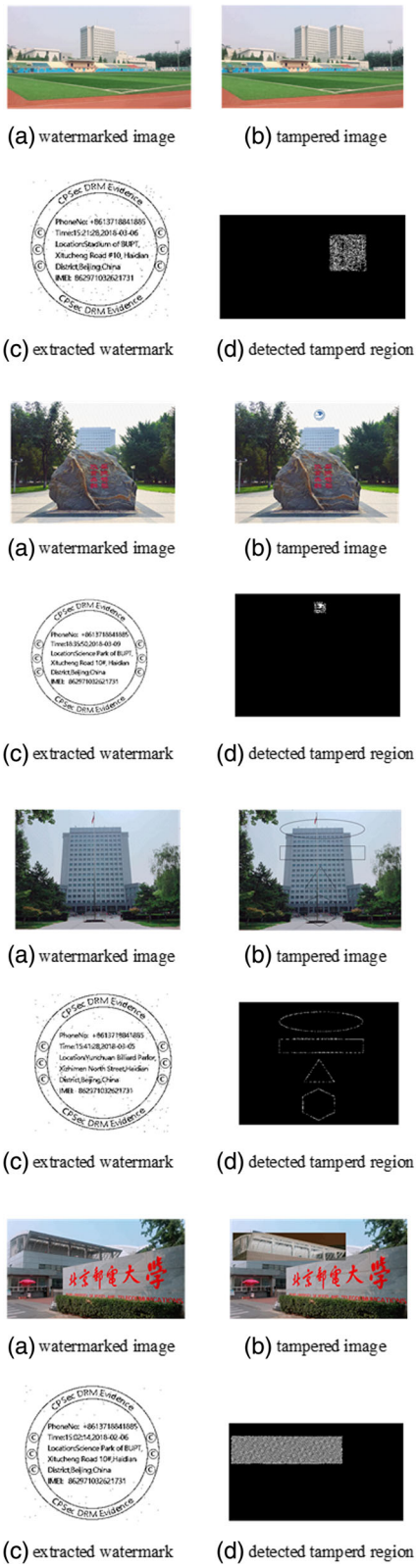


Fig. 12 Simulation results

Table 2 Experiments for different images when quality is 90

Image number	Image	CIR (%)	PSNR (dB)	Bit error (%)
1	Stadium	− 29.57	38.69	0.0001
2	Motto	− 29	39.92	0
3	Research Building	− 34.98	35.95	0.0001
4	Science park	− 29.67	38.98	0.0003

1. *Automatic watermark generation*: Compared with some other watermark scheme, the watermark image is generated automatically, using templates and photo information.
2. *Parameter control for watermark performance*: Jpeg compression can control the quality of watermarked mobile image, so as to balance the capacity of the image file and watermark performance.
3. *Scheme proposed misuse-tracing method based on robust watermark in post-usage stage*: The proposed watermark scheme supports robust watermark and semi-fragile watermark, which can certify mobile phone source authentication and detect the integrity and tamper tracking of the mobile pictures, even if the watermarked image was attacked by variant tamping.

Through experiments, we can see that the proposed watermark scheme is practical and effective, which can track source forensics and recognize the mobile-related information to trace and find the one who misuse or violate the confidential mobile pictures. As a highlight of the scheme, we can track where the attacks occurred.

In Table 3, we can get the watermark algorithm that is embedded randomly in the host image, which is security. The bit error is 0 even when the tampering ratio is up to 2.6%. And the algorithm is effective especially when color inverse attack happens; the bit error is very slow in experiment.

4 Conclusions

We proposed a new watermark-based mobile image data security scheme to protect confidential image data in this paper. We embedded some mobile-identification-related and user-identification-related information into the mobile photo after the mobile photos were taken, such as phone no., time, location, and IMEI of the mobile phone, which were robust watermarks for traceability and source authentication confirmation. Once the image data was misused such as spreading the confidential image in the Internet or for commercial purposes without authorization, we can identify and trace the misused responsibility by extracting the watermark embedded in the image data. Finally, we evaluated the proposed watermark scheme for a mobile by groups of variant size image data for security and efficiency, a large amount of groups of experiments manifest the proposed scheme was secure, efficient, pervasive, and robust for confident image data protection and misuse tracing.

Table 3 Experiments for tampering

Image number	Image	Tampering ratio (%)	Bit error (%)
1	Stadium	5.3	0.0039
2	Motto	1	0
3	Research building	1.4	0.0041
4	Science park	19	0.0015

Abbreviations

DCT: Discrete cosine transfer; DRM: Digital rights management; DWT: Discrete wavelet transfer

Acknowledgements

Not applicable.

Funding

This work is supported by the National Natural Science Fundamental of China (No. 61272519, No. 61170297, No. 61572080, No. 61472258).

Availability of data and materials

The datasets used is based on benchmark dataset Lenna, and the real data is based on the mobile Internet Phone with the type GIONEE M6.

Authors' contributions

The work is mainly finished by author MZ, who designed the watermark model and algorithm and wrote the whole paper. Author JM co-operated for the experiments and checking the references. Both authors read and approved the final manuscript.

Competing interests

The authors declare that they have no competing interests.

Publisher's Note

Author details

[1]Information Security Center of Beijing University of Posts and Telecommunications, Beijing 100876, China. [2]Audio and Video department, The Third Research Institute of China Electronics Technology Group Corporation, Beijing 100015, China.

Received: 19 May 2018 Accepted: 16 July 2018
Published online: 03 August 2018

References

1. A Piva, An overview on image forensics. ISRN Signal Processing **2013**, 1–22 (2013)
2. MC Stamm, M Wu, KJR Liu, Information forensics: an overview of the first decade. IEEE Access **1**, 167–200 (2013)
3. L Ozparlak, I Avcibas, Differentiating between images using wavelet-based transforms: a comparative study. IEEE Trans on Information Forensics and Security **6**(4), 1418–1431 (2011)
4. H Farid, MJ Bravo, Perceptual discrimination of computer generated and photographic faces. Digit. Investig. **8**(3/4), 226–235 (2012)
5. TY Chang, SC Tai, GS Lin, A passive multi-purpose scheme based on periodicity analysis of CFA artifacts for image forensics. J. Vis. Commun. Image Represent. **25**(6), 1289–1298 (2014)
6. P Fei, L Juan, L Min, Identification of natural images and computer generated graphics based on hybrid features. International Journal of Digital Crime and Forensics **4**(1), 1–16 (2012)
7. X Wang, L Yong, B Xu, et al., A statistical feature based approach to distinguish PRCG from photographs. Comput. Vis. Image Underst. **128**(11), 84–93 (2014)

8. P Fei, L Jiaoting, L Min, Identification of natural images and computer-generated graphics based on statistical and textural features. J. Forensic Sci. **60**(2), 435–443 (2015)
9. S Gao, Z Cong, C Wu, et al., *A hybrid feature based method for distinguishing computer graphics and photo-graphic image[G] //LNCS 8398:Proc of the Int Workshop on Digital-Forensics and Watermarking 2013* (Springer, Berlin, 2013), pp. 303–313
10. L Zhaohong, Z Zhenzhen, YQ Shi, Distinguishing computer graphics from photographic images using a multiresolution approach based on local binary patterns. Security and Communication Networks **7**(11), 2153–2159 (2014)
11. ALS Orozco, JR Corripio, LJG Villalba, et al., Image source acquisition identification of mobile devices based on the use of features. Multimedia Tools and Application **75**(12), 7087–7111 (2016)
12. C Chen, MC Stamm, *Camera model identification framework using an ensemble of demosaicing features // Proc of the Int Workshop on Information Forensics and Security* (IEEE, Piscataway, 2015), pp. 1–6
13. Z Deng, A Gijsenij, Z Jingyuan, *Source camera idenfication using auto-white balance approximation // Proc of the Int Conf ono Computer Vision* (IEEE, Piscataway, 2011), pp. 57–64
14. K Goyal, R Panwar, N Khanna, *Evaluation of IQM's Effectiveness for Cell Phone Identification Using Captured Videos and Images //Proc of the Int Conf on Power, Control and Embedded System* (IEEE, Piscataway, 2014), pp. 1–6
15. F Razzazi, Seyedadabi, *A robust feature for single image camera identification using local binary patterns // Proc of the IEEE Int Symp on Signal Processing and Information Technology* (IEEE, Piscataway, 2014), pp. 462–467
16. TH Thai, R Cogranne, F Retraint, Camera model identification based on the heteroscedastic noise model. IEEE Trans. Image Process. **23**(1), 250–263 (2014)
17. L Shuhan, X Kong, B Wang, et al., *Silhouette coefficient based approach on cell-phone classification for unknown source images //Proc of the 2012 IEEE Int Conf on Communications* (IEEE, Piscataway, 2012), pp. 6744–6747
18. J Fridrich, *Sensor Defects in Digital Image Forensic* (Springer, Berlin, 2013), pp. 179–218
19. AR Soobhany, KP Lam, P Fletcher, et al., *Mobile camera source identification with SVD // LNEE 313: Proc of the Int Joint Conf on Computer Information and Systems Sciences and Engineering* (Springer, Berlin, 2015), pp. 123–131
20. C Mo, J Fridrich, M Goljan, et al., *Source digital camcorder identification using sensor photo response non-uniformity //Proc of the 9th Conf on Security, Steganography, and Watermarking of Multimedia Contents IX* (SPIE, San Francisco, 2007) 65051G–65051G-12
21. G Hongmei, A Swaminathan, M Wu, Intrinsic sensor noise features for forensic analysis on scanners and scanned images. IEEE Trans on Information Forensics and Security **4**(3), 476–491 (2009)
22. B Mahdian, S Saic, Blind authentication using periodic properties of interpolation. IEEE Transactions of Information Forensics and Security **3**(3), 529–538 (2008)
23. AC Popescu, H Farid, Exposing digital forgeries by detecting traces of resampling. IEEE Transactioni on Signal Processing **53**(2), 758–767 (2005)
24. Z Zhen, K Jiquan, P Xijian, Blind detection of image splicing based on image quality metrics and moment features. Journal of Computer Applications. **28**(12), 3108–3111 (2008)
25. C Chen, YQ Shi, S Wei, in *Proc.of International Conference on Pattern Recognition*. A machine learning based scheme for double JPEG compression detection (United states: Institute of Electrical and Electronics Engineers, Tampa, 2008), pp. 1–11
26. W Luo, Z Qu, J Huang, in *Proc.of IEEE International Conference on Acoustics, Speech and Signal Processing*. A novel method for detecting cropped and recompressed image block (Institute of Electrical and Electronics Engineers, Honolulu, 2007), pp. 217–220
27. H Farid, Exposing digital forgeries from JPEG ghosts. IEEE Transactions on Information Forensics and Security **4**(1), 154–160 (2009)
28. S Bayram, I Avcibas, B Sankur, Image manipulation detection. Journal of Electronic Imaging **15**(4), 12–18 (2006)
29. J Fridrich, J Lukas, M Goljan, in *Proc.of the International Society for Optical Engineering*. Detecting digital image forgeries using sensor pattern noise (SPIE, San Jose, 2009), pp. 118–126
30. A Popescu, H Farid, Exposing digital forgeries in color filter array interpolated images[J]. IEEE Transactions on Signal Processing. **53**(10), 3948–3959 (2005)
31. F Di Martino, S Sessa, Fragile watermarking tamper detection with images compressed by fuzzy transform. Information Sciences **195**(13), 62–90 (2012)
32. S Rawat, B Raman, A chaotic system based fragile watermarking scheme for image tamper detection. AEU-International Journal of Electronics and Communications. **62**(10), 840–847 (2011)
33. RO Preda, Semi-fragile watermarking for image authentication with sensitive tamper localization in the wavelet domain. Measurement **46**(1), 367–373 (2013)
34. Z Lin, J He, X Tang, C-K Tang, Fast, automatic and fine-grained tampered JPEG image detection via DCT coefficient analysis. Pattern Recognition **42**(11), 2492–2501 (2009)
35. H-C Wu, C-C Chang, Detection and restoration of tampered JPEG compressed images. Journal of Systems and Software **64**(2), 151–161 (2002)