# Multiple Heterogeneous JPEG Image Hierarchical Forensic

**Xiangwei Kong, Bo Wang, Mingliang Yang and Yue Feng**

**Abstract** Since image processing software is widely used to tamper or embed data into JPEG images, the forensics of tampered JPEG images now plays a considerable important role. However, most existing forensics methods that use binary classification can hardly deal with multiclass image forensics problems properly under network environments. In this paper, we propose a hierarchical forensics scheme against multiple heterogeneous JPEG images. We introduce a compression identifier based on Markov model of DCT coefficients as the first hierarchical section and then develop a tampering detection and steganalyzer separately as the second phase. We conduct a series of experiments to testify the validity of the proposed method.

**Keywords** Image forensics · Heterogeneous images · Classification

## 1 Introduction

With the popularity of image editing tools such as Photoshop and steganography software, it becomes more convenient to modify digital images without leaving any perceptible artifacts. That makes digital image forensics become an increasingly heated issue.

X. Kong (✉) · B. Wang · M. Yang · Y. Feng
School of Information and Communication Engineering, Dalian University of Technology, Dalian 116024, People's Republic of China
e-mail: kongxw@dlut.edu.cn

B. Wang
e-mail: bowang@dlut.edu.cn

M. Yang
e-mail: yangml@mail.dlut.edu.cn

Y. Feng
e-mail: fy2012@mail.dlut.edu.cn



J.J. (Jong Hyuk) Park et al. (eds.), *Advanced Multimedia and Ubiquitous Engineering*, Lecture Notes in Electrical Engineering 393,
DOI 10.1007/978-981-10-1536-6_66

As illustrated in [1], binary-classification image forensics has made much progress [2] detected whether the part of an image was initially compressed at a lower quality than the rest [3–6] are put forward to distinguish single and double JPEG compression. Another branch of binary-classification image forensics is steganalysis forensics, a way to distinguish the original images and stego images that hide secret messages. The experimental results of [7] and [8] show that the detection accuracy attains above 90 % even for the low embedding rate of 10 %. However, images on network are not limited to binary-classification. So the existing image forensics algorithms cannot obtain high detection accuracy when the images are mainly from heterogeneous and multiclass source.

This paper focuses on the forensics of multiple heterogeneous images based on a hierarchical forensics scheme that aims at dealing with double compression,. followed by a tampering detection and steganalysis separately to forensics single compression level and double compression level.

## 2 The Proposed Hierarchical Forensics Scheme

From the view of forensics, JPEG images operation can be roughly divided into four classes. The first is the original JPEG images coming from camera or transformed from lossless images. The second is to scale or recompress and then resave in JPEG format. It happens when the images are uploaded to the social network. The above processed images can be considered as normally edited image whose content is not changed. The third is forged images whose content and information are changed. The forth is steganography that embeds data into images. These four classes can be further divided into binary classes. The original images have been compressed only once. As for the stego images, since the compression will destroy the secret messages, most steganographic algorithms are based on single compressed images. So they could also be regarded as single compressed. The other kind of images is edited or forged images. Double compression occurs when the image is originally stored in JPEG format and then resaved as a JPEG after tampering, because the second compression quality factor is different from the original one. Thus an image should be firstly detected whether is single or double compression.

As this hierarchical principle, we put forward the mechanism of hierarchical forensics. The first hierarchical layer is a compression identifier, which can identify the single compressed images and the double compressed images. The second layer is the forensics of single compressed images and double compressed image respectively.

## 3 Compression Identifier

As is proposed in [9], the first digits of DCT coefficients follow a generalized Benford's law for the single compression case, while for the double compression case, the distribution shows violation to the logarithmic trend. So the probability distribution of the first digits of DCT coefficients was used directly as features in [9] to classify double compressed images. However, the formed feature is first order statistics, which cannot reflect the correlations between adjacent DCT coefficients. So we model the distribution as a Markov chain and use one-step transition probability matrix to characterize this process. Being different from [9] and [10], we include 0 in the range of the first digits in order to retain information as much as possible. Matrix $F(i, j)$ represents the elements of the first digits of DCT coefficients in the position of $i$th row and $j$th column. The transition probability matrix for along the horizontal direction can be defined by:

$$p\{F(i,j+1) = v|F(i,j) = u\} = \frac{\sum_{i=1}^{m}\sum_{j=1}^{n-1} \delta(F(i,j)) = u, F(i,j+1) = v)}{\sum_{i=1}^{m}\sum_{j=1}^{n-1} \delta(F(i,j)) = u)}. \quad (1)$$

where $m$ and $n$ are the numbers of rows and columns respectively, and $\delta(x)$ equals one when $x$ is true, equals zero otherwise.

According to the stochastic processes theory, if a Markov chain has state space $S$ and transition probability matrix $p$, the stationary distribution $\pi$ will be unique when:

$$\lim_{n\to\infty} p^n(x, y) = \pi(y), y \in S. \quad (2)$$

In this case, $\pi$ can be calculated by solving the following equations:

$$\begin{cases} \pi = \pi p \ , \\ \sum_j \pi_j = 1 \ . \end{cases} \quad (3)$$

where $\pi$ is a 10-dimensional vector since there are 10 finite-states in the proposed Markov model. Since the stationary distribution $\pi$ of a Markov model is unique, it will be used as features for double compression detection in the proposed method.

As stated above, for a given image the DCT coefficients of the Y channel are extracted first in order to reduce the dimensionality. Then the first digits of DCT coefficients are calculated from the first 20 individual AC modes as stated in [10]. After that, each stationary distribution will be calculated according to (3) as features. As a result, a feature vector of $10 \times 20 = 200$ elements is obtained for each given image. And then the feature will be fed to the Support Vector Machine (SVM) classifier [11].

## 4 Tampering Detection and Localization

Figure 1 shows a classical image forged scenario. A region from a JPEG image (red lines) is pasted onto a host image (gray lines) and then recompressed in JPEG format (blue lines). The forged region usually exhibits the presence of non-aligned double JPEG (NA-JPEG) artifacts. As illustrated in Fig. 1b, the forged region is misaligned with the final JEPG compression block grid by shift ($x_f$, $y_f$), while the background region is misaligned with the final JEPG compression block grid by shift ($x_b$, $y_b$). Basing on the theory above, we can utilize the shift to locate the forged region.

When it comes to the relationship between shift and the DCT coefficients, let $m$ represent the number of 8 × 8 DCT blocks and $n(j)$ represent the sum of zero JPEG coefficients in the $j$th component. The percentage of zero JPEG coefficients in the $j$th component and the average percentage of zero JPEG coefficients are as follows:

$$p(j) = \frac{n(j)}{m}, j = 1, 2, \ldots, 64, \ \text{AVERAGE} = \frac{\sum_{1}^{64} p(j)}{64}. \tag{4}$$

To illustrate the relationship between AVERAGE and NA-JPEG compression shift, we crop a given JPEG image $I$ with quality factor QF along a block shift ($i$, $j$) and compress it a second time with the same quality factor to get image $I\_shift(i, j)$. We define the Double Block Shift Matrix (DBSM) of the given image $I$ as:

$$\boldsymbol{DBSM}(i,j) = \text{AVERAGE}(I\_shift(i,j)), \ 0 \le i,j \le 7 \tag{5}$$

We can find out that the position of the peak value in DBSM coincides with the JPEG block shift of NA-JPEG image. The given image is divided into overlapped image blocks in size of $N \times N$ with Step $M$. After that, we can use the shift ($i$, $j$) of every block to get the forged location of the given image.

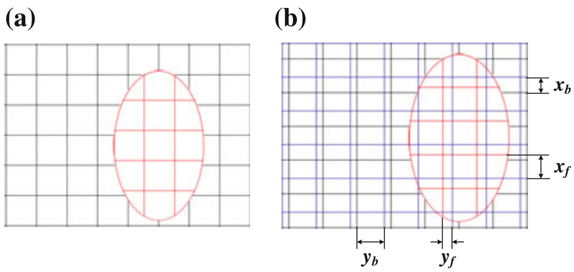


**Fig. 1** **a** Block artifacts before resaving; **b** block artifacts after resaving in JPEG format

# 5 Experiment and Discussion

We carried out our experiments on the BOSSbase database, which consists of 10,000 grayscale RAW images. We selected 500 images from randomly and compressed them in JPEG format with quality factor QF1 = {65, 75, 85, 95} respectively to generate 2000 single compressed images as the original JEPG images. The stego images were obtained by embedding a fixed relative payload of 0.1 bits per nonzero DCT coefficient (bpnc) with two steganographic algorithms, OutGuess [12] and F5 [13], generating 4000 stego images. Then each set of the original images are decompressed, cropped by 64 shifts $(i, j)$ $(0 \leq i, j \leq 7)$ separately, and compressed with QF2 = {65, 75, 85, 95, 99}. Finally we generated 640,000 double compressed JPEG images.

## *5.1 Compression Identifier Analysis*

We use SVM to classify the double and single compressed images. For each group, we randomly choose 50 % of the images as training data and the remaining for testing. To ensure the effectiveness and stability of the proposed method, the experiments are repeated 5 times with the training sets and testing sets chosen randomly. We made [10] as our comparison because in [1], it has been compared with other methods and it outperforms other previous methods such as [14] when it comes to distinguishing between single and double compression.

We can see from Table 1 that both [10] and our proposed method work well when QF2 > QF1. But for the situation of QF2 < QF1, our proposed method outperforms [10] in most cases. This is mainly because the artifacts can be magnified when considering the correlations between the neighbor DCT coefficients with Markov model.

**Table 1** Detection results of the compression identification (by %)

| QF2<br>QF1 | Method | 65 | 75 | 85 | 95 | 99 |
|---|---|---|---|---|---|---|
| 65 | Our method | – | 100 | 100 | 100 | 100 |
| | Work [10] | – | 97 | 97 | 98 | 99 |
| 75 | Our method | 99 | – | 99 | 100 | 100 |
| | Work [10] | 93 | – | 95 | 99 | 99 |
| 85 | Our method | 99 | 100 | – | 100 | 100 |
| | Work [10] | 89 | 89 | – | 96 | 98 |
| 95 | Our method | 95 | 97 | 100 | – | 100 |
| | Work [10] | 76 | 83 | 89 | – | 94 |

### 5.2 Tampering Location

For the cropped forged images, we compute the accuracy of each condition by averaging the accuracy of all the 64 shifts. The results are reported in Table 2. The accuracy in the case of QF2 > QF1 is much higher than in the case of QF2 $\leq$ QF1. This is because the block artifacts of the previous compression is weakened after the post compression when QF2 $\leq$ QF1.

### 5.3 Steganalysis

For JPEG images steganalysis, the CD-PEV feature vector [7] and the DCTR feature vector [15] are used to construct a SVM classifier. However, the performance of the method has a severe degradation for heterogeneous images. Table 3 shows the steganalysis will take the double compressed images as the stego images. So it is necessary to pick out the single compressed cover and stego images from the mixed images before the steganalysis classifier. Compared to binary classifier, the proposed method could successfully pick up the stego images from the mixed images.

**Table 2** Average accuracy (%) of the block shift estimate of the proposed method under different QF1 and QF2

| QF2<br>QF1 | 65 | 75 | 85 | 95 | 99 |
|---|---|---|---|---|---|
| 65 | 25.3 | 68.7 | 98.7 | 100 | 100 |
| 75 | 20.3 | 27.7 | 83.7 | 100 | 100 |
| 85 | 16 | 19 | 28.3 | 97.7 | 100 |
| 95 | 21 | 15.7 | 19.3 | 28 | 100 |

**Table 3** Detection accuracy (%) of steganalysis

| QF | Method | CD-PEV [7] | | DCTR [15] | |
|---|---|---|---|---|---|
| | | OutGuess | F5 | OutGuess | F5 |
| 65 | Our method | 93 | 90 | 95 | 97 |
| | Binary classifier | 50 | 50 | 49 | 50 |
| 75 | Our method\` | 95 | 92 | 96 | 96 |
| | Binary classifier | 50 | 50 | 50 | 50 |
| 85 | Our method | 96 | 94 | 96 | 96 |
| | Binary classifier | 50 | 49 | 50 | 47 |
| 95 | Our method | 97 | 97 | 96 | 96 |
| | Binary classifier | 47 | 47 | 50 | 50 |

## 6 Conclusion

In this paper, a scheme was proposed for the mixed images classification. We proposed to utilize the Markov model of DCT coefficients to identify the double compression. Experimental results show that our proposed method performs well even when QF2 < QF1. Then we presented an efficient method to locate the forged part in a tampered image. The proposed method does not need a classifier like a machine learning model and robust to common forgery processing such as resizing and blurring. After analyzing the double compression, the performance of the steganalysis can significantly outperform traditional steganalyzers in mixed images scenario.

**Acknowledgments** The work is supported by the Foundation for Innovative Research Groups of the NSFC(Grant No. 71421001), NSFC (Grant No. 61172109).

## References

1. Alessandro P (2013) An overview on image forensics. ISRN Sig Proc
2. Farid H (2009) Exposing digital forgeries from JPEG ghosts. IEEE Trans Inf Forensics Secur 4(1):154–160
3. Pevny T, Fridrich J (2008) Detection of double-compression in JPEG images for applications in steganography. IEEE Trans Inf Forensics Secur 3(2):247–258
4. Huang FJ, Huang JW, Shi YQ (2010) Detecting double JPEG compression with the same quantization matrix. IEEE Trans Inf Forensics Secur 5(4):848–856
5. Feng XY, Doerr G (2010) JPEG recompression detection. Proc SPIE 7541:75410J
6. Chen CH, Shi YQ, Su W (2008) A machine learning based scheme for double JPEG compression detection," In: 19th international conference on pattern recognition, pp 1–4
7. Pevný T, Fridrich J (2007) Merging Markov and DCT features for multi-class JPEG steganalysis. In: Proceedings SPIE security, steganography, and watermarking of multimedia contents, San Jose, CA, vol 6505, pp 1–13
8. Fridrich J (2005) Feature-based steganalysis for JPEG images and its implications for future design of steganographic schemes. In: Fridrich J (ed) Information hiding, 6th international workshop, volume 3200 of lecture notes in computer science, pp 67–81
9. Fu D, Shi YQ, Su W (2007) A generalized Benford's law for JPEG coefficients and its applications in image forensics. In: Proceedings of SPIE conference on electronic imaging, security and watermarking of multimedia contents, San Jose, USA
10. Li B, Shi YQ, Huang JW (2008) Detecting doubly compressed JPEG images by using mode based first digit features. In: IEEE international workshop on multimedia signal processing, pp 730–735
11. Chang CC, Lin CJ (2001) LIBSVM: a library for support vector machines, 2001. http://www.csie.ntu.edu.tw/~cjlin/libsvm
12. Provos N (2001) Defending against statistical steganalysis. In: 10th USENIX security symposium, Washington, D.C., pp 323–336
13. Westfeld A (2001) F5-a steganographic algorithm: high capacity despite better steganalysis. In: Proceedings of international workshop information hiding (IWIH), Pittsburgh, PA, pp 289–302

14. Chen YL, Hsu CT (2011) Detecting recompression of JPEG images via periodicity analysis of compression artifacts for tampering detection. IEEE Trans Inf Forensics Secur 6(2):396–406
15. Holub V, Fridrich J (2014) Low complexity features for JPEG steganalysis using undecimated DCT. IEEE Trans Inf Forensics Secur 10(2):219–28