Doctored JPEG Image Detection Based On Double Compression Features Analysis

Zhang Ting
CKC software lab, Ning Bo University
Ning Bo, 315211, China
wz_happy@126.com

Wang Rangding
CKC software lab, Ning Bo University
Ning Bo, 315211, China
wangrangding@nbu.edu.cn

Abstract—**Identifying the authenticity and integrity of digital images becomes increasingly important in digital forensics. In this paper, we focus on JPEG images and propose an effective method for detecting doctored images. We first investigate the statistical characteristics of DCT coefficients based on a recompression files sets, and analyze the differences of double compression effect between doctored and non-doctored region in a doctored image. We then extract the DCT coefficients histograms of each block in doctored images and represent them as feature vectors. We identify the location of doctored region by using SVM classification for evaluating the feature vectors. Experimental results demonstrate that the proposed method can efficiently detect and automatically locate doctored regions on different forgeries with low computational complexity.**

Keywords-Blind image forensics;image forgery;tempering detection;JPEG double compression

I. INTRODUCTION

Rapid advancement in image processing technology and software has made it remarkably easy to manipulate digital image without leaving any trace. This inevitably relates to some problems on legal forensics in digital image, such as digital image authentication, image medium copyright, individual privacy protection and so on. Proliferation of forgery photos will lead to terribly negative influence for society and personal life. Therefore, digital authentication has become an especially important research subject. Blind digital image forensics technology [1] which is a research hotspot in image processing authenticates doctored image depending on the statistical characteristics of media data itself without embedding any signature or watermarking beforehand. Farid developed several statistical methods for detecting forgeries based on region duplication [2], color filter interpolation [3], re-sampling [4], and lamp-house orientation [5]. Fridrich presented digital camera pattern noise to detect forgery [6-7]. Ng and Chang[8-9] proposed models of image spicing for detecting photomontage and physics-based model for distinguishing Computer Graphics from natural photographs. However, blind authentication studying is just in its baby-phase as many problems are pendent.

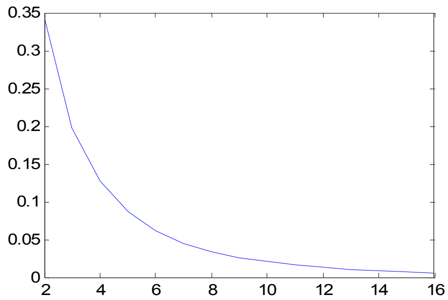
A common form of digital image tampering is removal, where region is copied and pasted to other part of the image for concealing important object or bringing inveracious information. Several researchers have developed methods for detecting this form of forgery. In [10], the authors analyzed the DCT coefficients for each block, while method [2] employs principal component analysis to capture the image blocks' features. But both the methods generated a large sort matrix which is the root cause of enormously computational complexity. Additionally, matching searching technique in [10], [2] is only fit to detect region duplication within non-compressed format image.

In this paper, we propose an effective method based on SVM multi-class classification for detecting doctored JPEG image. The algorithm regards area not possessing double compression features as suspected doctored region through analyzing recompression statistical characteristics. Further, doctored blocks are located by SVM and the doctored region consists of all the connected blocks. The experimental results show that the proposed method can successfully detect copy-paste forgery within or between JPEG images with low computational complexity.

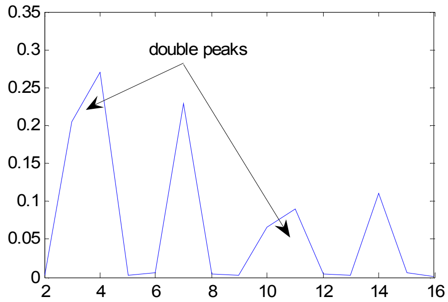
II. DOUBLE JPEG COMPRESSION ANALYSIS

A. *Double JPEG Compression Features*

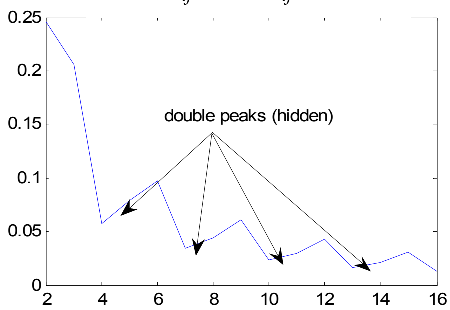
Double JPEG compression is the JPEG secondary compression process for image with different quantization matrix Q^1 (initial matrix), Q^2 (secondary matrix). If and only if $Q^1 \neq Q^2$, the DCT coefficients D_{ij} are considered as being doubly compressed. Before the secondary compression, all the AC coefficients D_{ij} histograms obey Laplace distribution. And when recompressed, they form local maximum, minimum and double peaks features in D_{ij}^q histogram. For instance, Fig 1 shows that the DCT coefficients D_{ij}^q probability histograms, in which standard quantization tables are used (ideal case). In the Fig 1, lateral axis expresses the multiples of secondary quantization factor but multiple 0 and 1, where the probability values are too big to have distinguishable characteristics. Vertical axis expresses the normalized statistical values. When $Q_{ij}^1 = Q_{ij}^2$, the histogram has no any feature and the curve is flat as shown in Fig 1(a). When $Q_{ij}^1 > Q_{ij}^2$, in Fig 1(b), some multiples form obvious double peaks and some unquantized multiples form local minimum 0. In Fig 1(c), all the 0 value points contribution to double peaks which appear between two local maximum values, then, 0 value points disappear and the double peaks are invisible. When $Q_{ij}^1 < Q_{ij}^2$ and Q_{ij}^1 is not a common factor of Q_{ij}^2, in Fig 1(d), the double peaks appear but are invisible after local maximum values.



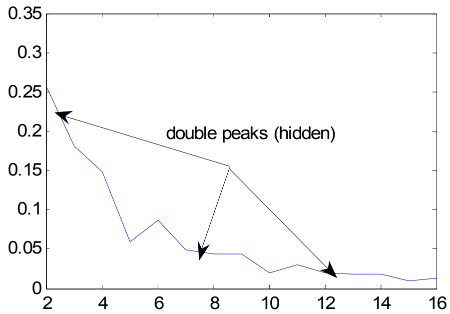
(a) $Q_{ij}^1 = 5, Q_{ij}^2 = 5$



(b) $Q_{ij}^1 = 7, Q_{ij}^2 = 2$



(c) $Q_{ij}^1 = 6, Q_{ij}^2 = 4$



(d) $Q_{ij}^1 = 5, Q_{ij}^2 = 8$

Fig. 1 Normalized DCT coefficients histograms (ideal case)

According to Fig 1, we can see that different combinations of initial quantization factor (IQF) Q_{ij}^1 and the secondary quantization factor (SQF) Q_{ij}^2 make histograms behave differently. Thereby, the features can be used to estimate IQF Q_{ij}^1 under the condition of the same Q_{ij}^2. However, D_{ij}^q histogram may be undistinguishable in some conditions which make IQF Q_{ij}^1 fail to be estimated correctly. Such as, when $Q_{ij}^1 = 1$, $Q_{ij}^1 = Q_{ij}^2$ and Q_{ij}^1 is a common factor of Q_{ij}^2, their histograms have no any feature and the curves are similar with Fig 1(a). Here, the Q_{ij}^1 can not be distinguished in the three situations.

B. Double Compression Effect in Doctored JPEG Image

A doctored JPEG image is usually compressed secondarily. The secondarily compressed image includes doctored and non-doctored region. The non-doctored region exhibits double compression feature as corresponding region of originally double compression image. The doctored region either from non-compression image (BMP format), or from original image itself or other JPEG image, exhibits no double compression feature. The process of copy-paste is given in Fig 2, in which the upper layer is doctored cut-paste region and the bottom layer is original background area. When pasted into original image, the DCT blocks may take place position displacements, which cause the pasted DCT blocks to be divided into four sub blocks respectively falling into adjacent four blocks of original image. Thus, the DCT coefficients of four sub blocks are irreversible losing double compression features because of irreversibility of block DCT and IDCT in the compression and decompression process. Moreover, the DCT blocks including doctored edge also exhibit no double compression and some operations to doctored edge, such as, smoothing, feathering etc, as well as make them lose double compression feature.

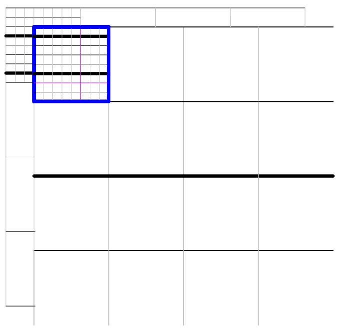


Fig. 2 position shifts of DCT blocks in a doctored part

III. FORGERY DETECTION ALGORITHM

A. Multi-class SVM Based on Initial Quantization Factor Estimation

The IQF can be estimated through quantified DCT coefficients histograms features in large numbers of double compression images. These histograms will be divided into the same class with the same IQF and SQF. Therefore, the estimation to IQF will be the classification problem about DCT coefficients histograms.

In this paper, we adopt One-Against-One recognition algorithm in SVM multi-class classification. Give k classes data, a classifier is designed via selecting arbitrary i^th and j^th class data, where $i < j$ ("i" is positive example, and "j" is negative example). Thus, $k(k-1)/2$ numbers of classifiers require to be trained. To i^th and j^th class data, given t groups data $(x_i, y_i),...,(x_t, y_t)$, $i = 1,...,t$, $x_i \in R^{n}$, $y \in \{+1, -1\}^t$, the following problems need to be solved.

$$\min_{\mathbf{w^{ij}}, b^{ij}, \xi_t^{ij}} \frac{1}{2}(\mathbf{w^{ij}})^{\mathrm{T}} \mathbf{w^{ij}} + C\sum_t \xi_t^{ij} (\mathbf{w^{ij}})^{\mathrm{T}} \quad (1)$$

$$(\mathbf{w^{ij}})^{\mathrm{T}} \phi(x_i) + b^{ij} \geq 1 - \xi_t^{ij} \quad (2)$$

$$(\mathbf{w^{ij}})^{\mathrm{T}} \phi(x_i) + b^{ij} \leq -1 + \xi_t^{ij} \quad (3)$$

Where, ξ_t^{ij} is non-relaxation variable introduced by training samples set when they are linearly inseparable. The $\mathbf{w^{ij}}$ is weighted support vector and b^{ij} is parametric variable. Function $\phi(x)$ transforms $\mathbf{x_i}$ to a higher-dimensional (or infinite-dimensional) linear space. The equilibrium values $\mathbf{w^{ij}}$, b^{ij} , ξ_t^{ij} between regularization term $(\mathbf{w^{ij}})^{\mathrm{T}} \mathbf{w^{ij}} / 2$ and training error are computed by equation (1), in which minimizing $(\mathbf{w^{ij}})^{\mathrm{T}} \mathbf{w^{ij}} / 2$ is equal to maximize $2 / \left\| \mathbf{w^{ij}} \right\|$, namely, maximal interval between two groups data. When linearly inseparable, the penalty factor $C\sum_t \xi_t^{ij}$ can be used to decrease training error value. Then, compute $y_t((\mathbf{w^{ij}})^{\mathrm{T}} \phi(x_i) + b^{ij})$ value and judge constraint conditions. If equation (2) is fitted, x belongs to i^th class, else if equation (3) is fitted, x belongs to j^th class. Otherwise, when $y_t((\mathbf{w^{ij}})^{\mathrm{T}} \phi(x_i) + b^{ij}) = 0, \xi_t^{ij} = 1$, x is the class of smaller index value.

B. Detection Algorithm

After DCT transformed, the image energy is concentrated on low frequency area and DCT coefficients in this area will be more sensitive than others when the image undergoes doctoring, so we select them to construct features set. Repeated experiments demonstrate that the D_{10}^{th} DCT coefficient in each DCT block, exhibits abundant variation characteristics in the recompressed image. According to the differences of double compression features between doctored and non-doctored region, the algorithm first constructs the DCT coefficient D_{10}^{q} histograms in recompressed image, then, it uses SVM to classify constructed feature vectors. The concrete steps are as follows:

Step1: A sets of BMP image samples are doubly JPEG compressed, and the compression quality(quantization factor) are 61(9),65(8),70(7),75(6),78(5),82(4),86(3),92(2),95(1). For each specific secondary compression quality, the images with initial compression quality 61~95 are the same class. The sample sets are classified 9 classes totally.

Step2: The D_{10}^{q} histogram of each double compression file is computed and formed a feature vector $\{h(2), h(3),..., h(16)\}$, where $h(m)$ is the statistical value of m times Q_{ij}^2 .

Step 3: SVM classifier design: RBF is the kernel function. The input is all categories of feature vector sets in each class, and the output is the category of each feature vector. For each class data, a part of them are used to train and the remaining are used to test. In the training process, all the outputs make up of a training model file.

Step 4: The information in JPEG head file is extracted and the D_{10}^{th} position quantization factor in luminance quantization table is obtained.

Step 5: Divide the image into non-overlapping N sub-blocks from left upper to right bottom corner and number them orderly, where N is equal to 16~81.

Step 6: compute the histogram of D_{10}^{q} coefficients in each block and construct it as a feature vector.

Step 7: Combining with quantization factor, load the training model file obtained in Step3 and predict the category of each block orderly.

Step 8: mark blocks with identical category the same color by the classification results then the doctored and non-doctored region are distinguished.

IV. EXPERIMENTS AND RESULTS ANALYZING

One of image forgery technologies is copy-paste. For this form of forgery, we utilize the proposed algorithm to detect several tempered images of daily life. In this experiment, the standard quantization tables are used in compressing process, and the related parameters are Q_{ij}^1, Q_{ij}^2 and number of sub-block N. The following experiments verify the detected effects using our method in several forgeries with different double compression quantization factor combinations.

A BMP background image, size of 3028X2271, is JPEG compressed with quality factor (quantization factor) 78(5). Another BMP portrait is compressed with 70(7). The compressed images are copied and pasted to generate a forgery image, shown in Fig 3. Then, the forgery is JPEG compressed with quality factor 92(2), 82(4) and 65(8) respectively.

The Fig 4 shows the detected results of Fig 3(c). In the Fig 4, the black part is original image area possessing double compression features and the algorithm can estimate its IQF. The white part which lies in above of the image is the sky areas with single details. This belongs to error detection situation because of lacking sufficient statistical characteristics after recompression, and the algorithm can not correctly estimate its IQF. The part of people which possesses rich details and good statistical characteristics is the tempered area. The Fig 4(a),(b),(c) display detection results when the related parameters are: number of sub-block N=16, background IQF $Q_{ij}^1 = 5$, SQF Q_{ij}^2 =2(best), Q_{ij}^2 =4(high), Q_{ij}^2 =8(middle). As observed, the detection effects gradually become poor with the increase in SQF. The Fig4 (d) shows the detection result when N=49, background IQF $Q_{ij}^1 = 5$, SQF $Q_{ij}^2 = 2$. Compared with Fig4 (a), the detection effect gradually becomes well with increasing block number under the condition of same SQF.

In order to measure the detection effect when the background image has different IQF, we replace the background with a richly detailed image and make a forgery according to the above same method. In the Fig 5(c), the IQF of background Q_{ij}^1 is 61(9) ~92(2) and SQF Q_{ij}^2 is 92(2) (best). Thus, the 8 tempered images are produced with different combinations of Q_{ij}^1, Q_{ij}^2.

The detection results of all forgeries generated in Fig 5 are displayed in Fig 6.The number of sub-block in the tempered images is N=81. As seen, the detection results among Fig 6(a) ~ (g) have slight differences but they all can successfully locate the tempered area. Sporadic blocks in Fig 6(d), (g) are wrongly detected however they can be eliminated based on the analysis in Fig 4. While, in the Fig 6(h), the tempered area can not be detected successfully at all because initial compression and secondary compression quality in background are equal, that is, $Q_{ij}^1 = Q_{ij}^2$ then, both background and tempered area exhibit similarly single compression phenomenon.

When one more regions are tempered, the manipulation may be within or between images. So we produce a copy-paste forgery containing three positions of tempering, shown in Fig 7(b). Here, the two tempered regions are within the image itself and the one is from else JPEG image. The number of sub-block is N=64 and IQF, SQF are respectively Q_{ij}^1 =78(5), Q_{ij}^2 =92(2). Figure 7(c) is the detected result, in which all the tempered areas are detected successfully except wrongly detected phenomenon in image areas with similar texture.

We can see from the above experiments, our method can work in detecting copy-paste forgeries effectively. For some false alarms, they can be eliminated by visible judgment. The detection effect may be improved to some extent in reasonably dividing sub-block and increasing number of sub-block. Generally, for other double compression combinations of Q_{ij}^1, Q_{ij}^2 outside our experimental situations, the algorithm can also detect them effectively. Furthermore, the paper extends conventional methods of detecting region duplication within an image to detect copy-paste between images effectively.

V. CONCLUSION

The proposed forgery detection method based on double compression features can identify copy-paste tempering for JPEG images. For other forgery technologies, the algorithm can detect validly so long as the forger has undergone double compression process. But when the size of tempered area is quite tiny, the algorithm may fail. This algorithm is also limited for detecting some images with single details and similar texture. Consequently, the future work is to explore new scheme adapting to background image with extremely single details.

REFERENCES

[1] Ng T T, Chang S F, Lin C Y, and Qibin Sun Q B, "Passive-Blind Image Forensics," Multimedia Security Technologies for Digital Rights, Zeng W, Yu H, Lin C Y(eds.), Elsvier, 2006.

[2] Popescu A C , Farid H, "Exposing digital forgeries by detecting duplicated image regions,"Dartmouth College , Hanover ,New Hampshire , USA, TR2004-515 , 2004.

[3] Popescu A C , Farid H, "Exposing digital forgeries in color filter array interpolated images,"IEEE Transactions on Signal Processing , pp. 3948-3959, 2005.

[4] Popescu A C , Farid H, "Exposing digital forgeries by detecting traces of resampling,"IEEE Transactions on Signal Processing , pp. 758-767, 2005.

[5] Johnson M K, Farid H, "Exposing digital forgeries by detecting inconsistencies in lighting,"Proceedings of the 7th Workshop on Multimedia and Security, New York, NY,USA, pp. 1-10, 2005.

[6] Lucas J , Fridrich J , and Goljan M, "Digital " bullet scratches" for images," Proceedings of the IEEE International Conference on Image processing. Genova , Italy , pp. 65-68, March 2005.

[7] Lukcas J , Fridrich J , and Goljan M, "Detecting digital image forgeries using sensor pattern noise,"Proceedings of the SPIE :Security , Steganography , and Watermarking of Multimedia Contents. San Jose , California,USA, pp. 362-372, 2006.

[8] Ng TT , Chang S F, "A model for image splicing," Proceedings of the IEEE International Conference on Image Processing, Singapore , pp. 1169-1172, 2004.

[9] Ng T T , Chang S F , and Hsu J et al, "Physics-motivated features for distinguishing photographic images and computer graphics,"Proceedings of the ACM Multimedia. Singapore , pp. 239-248, 2005.

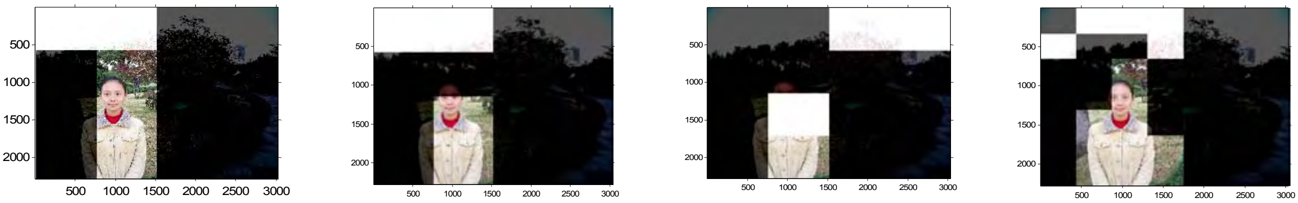
[10] Fridrich J , Soukal D , and Lukas J, "Detection of copy-move forgery in digital images," Proceedings of the Digital Forensic Research Workshop. Cleveland OH , USA , 2003.

[11] Fan Z, de Queiroz R. L, "Identification of Bitmap Compression History: JPEG Detection and Quantizer Estimation," in IEEE Transactions on Image Processing, vol. 12, pp. 230–235, February 2003.

[12] Chih-Wei Hsu, Chih-Jen Lin, "A comparison of methods for multiclass support vector machines," Neural Networks, IEEE Transactions on, vol. 13, pp. 415-425, 2002.

(a) The original background (b) The original portrait (c) The forgery

Fig. 3 Test images

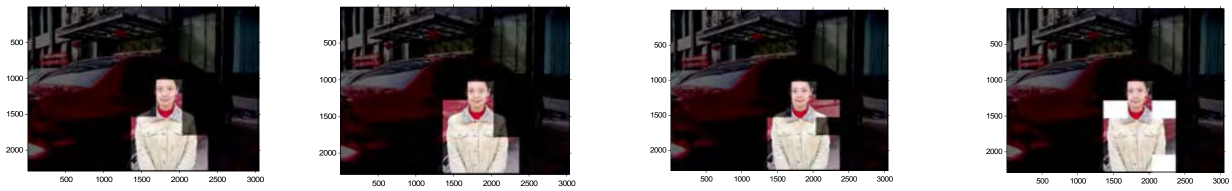


（a） $Q_{ij}^1 = 5, Q_{ij}^2 = 2$ （b） $Q_{ij}^1 = 5, Q_{ij}^2 = 4$ （c） $Q_{ij}^1 = 5, Q_{ij}^2 = 8$ （d） $Q_{ij}^1 = 5, Q_{ij}^2 = 2$

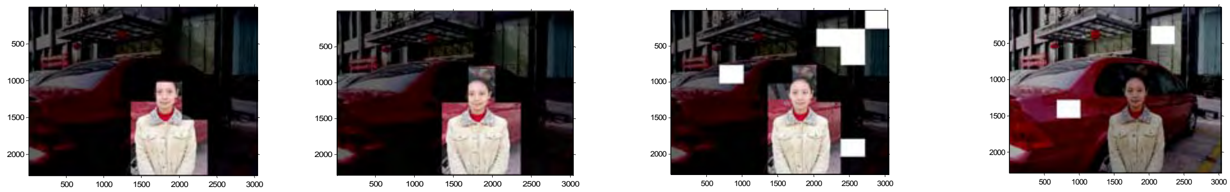
Fig. 4 Detection results of doctored images in Fig 3

(a) The original background (b) The original portrait (c) The forgery

Fig. 5 Test images



(a) $Q_{ij}^1 = 9$, $Q_{ij}^2 = 2$ (b) $Q_{ij}^1 = 8$, $Q_{ij}^2 = 2$ (c) $Q_{ij}^1 = 7, Q_{ij}^2 = 2$ (d) $Q_{ij}^1 = 6, Q_{ij}^2 = 2$



(e) $Q_{ij}^1 = 5$, $Q_{ij}^2 = 2$ (f) $Q_{ij}^1 = 4$, $Q_{ij}^2 = 2$ (g) $Q_{ij}^1 = 3$, $Q_{ij}^2 = 2$ (h) $Q_{ij}^1 = 2, Q_{ij}^2 = 2$

Fig. 6 Detection results of doctored images in Fig5

(a) The original image (b) The forgery (c) The detection result

Fig. 7 The original, doctored images and detection result