Detection of Shifted Double JPEG Compression using Markovian Transition Probability Matrix

Yujin Zhang, Shenghong Li and Shilin Wang
Shanghai Jiao Tong University, Shanghai
E-mail: {yjzhang82, shli, wsl}@sjtu.edu.cn Tel: +86-21-34205281

***Abstract*—Copy-paste forgery is a very common type of forgery in JPEG images. The tampered patch has always suffered from JPEG compression twice with inconsistent block segmentation. This phenomenon in JPEG image forgeries is called the shifted double JPEG (SD-JPEG) compression. Detection of SD-JPEG compressed blocks can make crucial contribution to detect and locate the tampered region. However, the existing SD-JPEG compression detection methods cannot achieve satisfactory results especially when the image block size is small. In this paper, an effective SD-JPEG compression detection method based on Markovian transition probability matrix is proposed. Statistical artifacts are left by the SD-JPEG compression among the elements of the JPEG 2-D arrays. Difference JPEG 2-D arrays generated along four directions (i.e. horizontal, vertical, main diagonal and minor diagonal) are utilized to enhance them and then thresholded by a predefined threshold for reducing computational cost. Markovian transition probability matrix is used to model the difference JPEG 2-D arrays in order to utilize the second order statistics. All the elements of these transition probability matrices are served as features for SD-JPEG compression detection. Support Vector Machine (SVM) is employed as the classifier. Experimental results demonstrate the efficiency of the proposed method.**

I. INTRODUCTION

With the rapid growth of cheap and high resolution digital cameras, high performance computers, and powerful image processing software, it is getting easier to manipulate a digital image without leaving obvious visual traces. As a matter of fact, the problem of digital image counterfeiting is potentially serious, sometimes malicious tampering may bring us some legal crisis. Therefore, digital image forensics has become an especially important research subject.

Generally speaking, existing approaches for digital image forensics can be divided into two categories: active [1, 2] and passive [3-5] approaches. Active approaches mainly insert watermarks or signatures to digital images at the time of recording, the authenticity of digital images can be checked by detecting the change in the watermark, however most image capturing devices don't equip with watermarking functionality due to the cost. In contrast to active approaches, passive approaches don't need any watermark or prior information about image, so they have drawn a lot of attention and become a hot research topic in digital image forensics.

It is well-known that JPEG is a commonly used image compression standard. Identifying the authenticity of JPEG format images plays a useful role in image forensics. Recently, several methods for JPEG image authentication based on double compression detection were proposed, the block segmentations of the first and second JPEG compressions are aligned to each other under these circumstances. Lukáš and Fridrich [6] tried to identify double JPEG compression by detecting DCT histogram artifacts of individual coefficients and estimate the primary quantization matrix from a double compressed image. Popescu and Farid [7] also pointed out that double quantization introduces periodic artifacts to the JPEG mode's histogram, and devised a quantitative measure for these artifacts to discriminate between single and double JPEG compressed images. Fu et al. [8] presented a novel statistical model based on Benford's law for the probability distributions of the first digits of the block-DCT and quantized JPEG coefficients and claimed that double JPEG compression could be detected because it would cause severe violation of the first digit law. Li et al. [9] observed that using generalized Benford's law to fit distribution of the first digits of JPEG coefficients from some selected individual AC modes, the performance of detecting double JPEG compression could be greatly enhanced, and then the quality factor in the primary JPEG compression could be identified by a multi-class classification strategy. Chen et al. [10] used Markov based transition probability matrix as the feature and employed the support vector machine as the classifier to distinguish double compressed images from single JPEG compressed images, the author stated that their algorithm outperformed the prior work in [7].

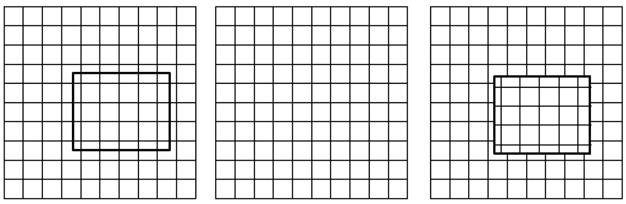
Copy-paste forgery is a very common type of forgery in JPEG images. The tampered region has always suffered from the JPEG compression twice with inconsistent block segmentation. In such situations, methods based on double compression don't work effectively. Luo et al. [11] developed a blocking artifact characteristics matrix (BACM) to detect SD-JPEG compression. They observed that the BACM exhibits regular symmetrical shape for the single JPEG compression, but symmetry property is destroyed after SD-JPEG compression. Chen and Hsu [12] presented a linearly dependency model of pixel differences, and analyzed the different peak energy distribution to discriminate SD-JPEG compressed images from single JPEG compressed images. Qu et al. [13] formulated the SD-JPEG compression as noisy convolutive mixing model, tried to solve it using blind signal separation. Rather than using ICA, the authors analyzed the Independent Value Map (IVM) to determine whether the image has been SD-JPEG compressed with the assumption that SD-JPEG compression will break the symmetry of IVM. However, these methods can achieve good performance only when the image block size is large, it is hard to employ them to detect and locate the some small tampered regions for composite JPEG images.

Motivated from the above discussion, in this paper, we first propose an effective SD-JPEG compression detection method based on Markovian transition probability matrix to identify the SD-JPEG compressed blocks, and then we use the proposed method to detect and locate the tampered regions for composite JPEG images.

The rest of this paper is organized as follows. The process of copy-paste forgery in JPEG images is introduced in section II. The proposed approach is described in section III. The experimental results with comparison to two existing methods are shown in section IV. Conclusion is made in section V.

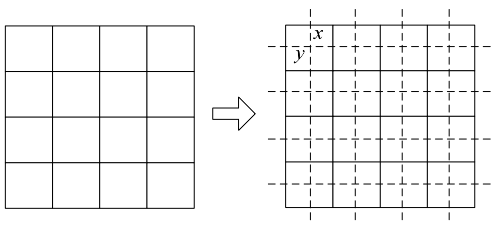
II. Model of Copy-paste Forgery in Jpeg Images

As we known, JPEG compression is a popular compression scheme based on block discrete cosine transform, the block artifacts can be as an inherent signature for JPEG compressed images. Copy-paste forgery is a very common type of forgery in JPEG images. It means that a patch of the source image is cropped and pasted it onto the target image to generate a new composite image, as illustrated in Fig.1.



(a) Original image1 (b) Original image 2 (c) Composite image

Fig.1. Demonstration of copy-paste forgery in JPEG images.



(a) (b)

Fig.2. Generation of a SD-JPEG compressed block (a) JPEG compressed with the quality factor QF_1 (b) SD-JPEG compressed with the quality factor QF_2 and block segmentation in dotted line with a coordinate shift (x, y).

In Fig.1, suppose that each grid cell represents a pixel block of size 8×8, the original image 1 is compressed with the quality factor QF_1, the original image 2 is uncompressed or compressed with the quality factor QF_2. The rectangle region is cropped from the original image 1, and then pasted onto the original image 2. The composite image is stored in JPEG format with the quality factor QF_2. We can notice that the blocking artifact grids of the tampered region are misaligned with the original ones with a probability of 63/64.Therefore, it is very reasonable to assume that the tampered region has undergone the SD-JPEG compression. Generation of a SD-JPEG compressed block is shown in Fig.2 so as to better illustrate the SD-JPEG compression.

In order to detect and locate the tampered regions of a composite JPEG image, the composite JPEG image can be segmented into small blocks and identified one by one.

III. Proposed Approach

In general, for each pixel of a color JPEG image, there are three numerical values that collectively describe its color. They are identified as Y, C_b, C_r. Y is the luma value, and C_b and C_r collectively form the chrominance value. Chrominance subsampling makes SD-JPEG compression detection in chroma spaces very difficult, hence we only consider the Y component. Fig.3 shows the framework of the proposed feature extraction procedure.

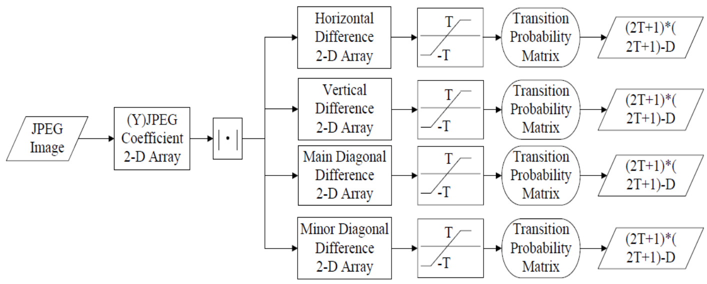


Fig.3. Diagram of the proposed feature extraction.

A. JPEG 2-D Array

For a given JPEG image, JPEG coefficient 2-D array is made up of all the quantized 8×8 block discrete cosine transform (BDCT) coefficients. The proposed approach only considers the magnitude elements of the JPEG coefficient 2-D array (JPEG 2-D array for short). Because the SD-JPEG compression leaves statistical artifacts among them. Whereas the sign of few quantized BDCT coefficients changes after SD-JPEG compression.

B. Difference JPEG 2-D Array

Elements in the JPEG 2-D array are often correlated for a single JPEG compressed image. The SD-JPEG compression artifacts disturb the JPEG 2-D array, weakening the correlation among the elements of the JPEG 2-D array. The SD-JPEG compression detection can be considered as a problem of weak signal detection in the background of strong signal. In order to reduce the effects caused by the diversity of image content and enhance the SD-JPEG compression artifacts, difference JPEG 2-D arrays are generated along four directions (i.e. horizontal, vertical, main diagonal and minor diagonal) as shown in (1).

$$\begin{aligned} E_h(i,j) &= x(i,j) - x(i+1,j) \\ E_v(i,j) &= x(i,j) - x(i,j+1) \\ E_d(i,j) &= x(i,j) - x(i+1,j+1) \\ E_m(i,j) &= x(i+1,j) - x(i,j+1) \end{aligned} \qquad (1)$$

Where $x(i, j)$ ($i \in [0, M\text{-}1]$, $j \in [0, N\text{-}1]$) indicates the JPEG 2-D array of a given JPEG image. M and N are the size of the JPEG 2-D array along the horizontal and vertical directions. The symbols E_h (i, j), E_v (i, j), E_d (i, j), E_m (i, j) denote difference JPEG 2-D arrays along horizontal, vertical, main diagonal and minor diagonal directions, respectively.

C. Thresholded Difference JPEG 2-D Array

A predefined threshold T is used to deal with the difference JPEG 2-D array as shown in (2). In doing so, the value range of the difference JPEG 2-D arrays is limited to [-T, T], with only 2*T+1 possible values.

$$E(i,j)=\begin{cases}E(i,j) & |E(i,j)|<T\\ T & E(i,j)\geq T\\ -T & E(i,j)\leq -T\end{cases} \tag{2}$$

The threshold T cannot be too small or too large. With a too small T, the transition probability matrix will not be able to sensitively catch the artifacts caused by SD-JPEG compression. With a too larger T, dimensionality of the transition probability matrix will be too high. Besides, more information about image content itself rather than SD-JPEG compression may be introduced with the increase of the threshold T. Based on the experimental dataset prepared in section IV, this threshold is selected as 2 in the proposed approach.

D. Transition Probability Matrix of Thresholded Difference JPEG 2-D Array

We model the thresholded difference JPEG 2-D arrays defined above by using one-step Markov random process. According to random process theory, a transition probability matrix (TPM) can be used to characterize a Markov random process. Equations (3) show the transition probability matrix for horizontal, vertical, main diagonal and minor diagonal difference JPEG 2-D arrays respectively. Consequently, we have 5×5 elements for the transition probability matrix along each direction. There are 100 elements in the feature vector for a given JPEG image. All the elements of these transition probability matrices for E_h, E_v, E_d and E_m are served as features for the SD-JPEG compression detection.

$$\begin{aligned}
p\{E_h(i+1,j)=n|E_h(i,j)=m\}&=\frac{\sum_{i=0}^{M-2}\sum_{j=0}^{N-2}\delta(E_h(i,j)=m,E_h(i+1,j)=n)}{\sum_{i=0}^{M-2}\sum_{j=0}^{N-2}\delta(E_h(i,j)=m)}\\
p\{E_v(i,j+1)=n|E_v(i,j)=m\}&=\frac{\sum_{i=0}^{M-2}\sum_{j=0}^{N-2}\delta(E_v(i,j)=m,E_v(i,j+1)=n)}{\sum_{i=0}^{M-2}\sum_{j=0}^{N-2}\delta(E_v(i,j)=m)}\\
p\{E_d(i+1,j+1)=n|E_d(i,j)=m\}&=\frac{\sum_{i=0}^{M-2}\sum_{j=0}^{N-2}\delta(E_d(i,j)=m,E_d(i+1,j+1)=n)}{\sum_{i=0}^{M-2}\sum_{j=0}^{N-2}\delta(E_d(i,j)=m)}\\
p\{E_m(i,j+1)=n|E_m(i+1,j)=m\}&=\frac{\sum_{i=0}^{M-2}\sum_{j=0}^{N-2}\delta(E_m(i+1,j)=m,E_m(i,j+1)=n)}{\sum_{i=0}^{M-2}\sum_{j=0}^{N-2}\delta(E_m(i+1,j)=m)}
\end{aligned} \tag{3}$$

Where $m, n \in Z$, $-T \leq m, n \leq T$, and

$$\delta(A=m,B=n)=\begin{cases}1, & A=m \text{ and } B=n\\ 0, & otherwise\end{cases} \tag{4}$$

IV. Experiments and Results

In our experiments, the Matlab JPEG Toolbox [14] is employed for JPEG compression. Images collected in our experiments are obtained from the UCID [15] and NRCS [16] image databases. All the original images in above databases are uncompressed. LIBSVM [17] is used as the classifier. The polynomial kernel with degree three for the proposed method and Radial Basis Function kernel for two existing SD-JPEG detection methods are selected.

A. Detection of the SD-JPEG Compressed Blocks

In the SD-JPEG compressed blocks detection experiment, the experimental dataset is created as follows:

(1) Three sizes (64×64, 128×128 and 256×256) of image blocks are generated for assessment.

(2) QF_1 and QF_2 are selected from 50 to 90 with an increment of 10.

(3) In a specified test condition, 1000 uncompressed image blocks are randomly extracted from the UCID database. For a given $\{QF_1, QF_2\}$ pair, there are totally 1000 pairs of SD-JPEG compressed blocks and single JPEG compressed blocks. Each pair of SD-JPEG compressed block and single JPEG compressed block have the same size and image contents for each bmp images. SD-JPEG compressed blocks are generated from the extracted blocks with the quality factor QF_1, QF_2 and random coordinate shifts. Single JPEG compressed blocks are generated from the same blocks only with the quality factor QF_2 and same coordinate shifts.

Therefore, for a given $\{QF_1, QF_2\}$ pair, there are totally 2000 image blocks in this dataset. Half of them, i.e. 500 SD-JPEG compressed image blocks and 500 single JPEG compressed blocks are randomly picked out to train the SVM classifier and others for testing. Grid searching is employed to select the best parameters for the classifier. 5-fold cross validation is used in classification. Detection accuracies are average over 20 random experiments.

To test the efficacy of the proposed method, two existing SD-JPEG detection methods, i.e., the BACM-based [11] and ICA-based [13] methods, have also been investigated for comparison on the same dataset. The detection accuracies of the BACM-based, ICA-based and proposed methods for block sizes of 64×64, 128×128 and 256×256 are shown in Table I, Table II and Table III respectively.

TABLE I DETECTION ACCURACY (%) WITH VARIOUS VALUES OF QF_1 AND QF_2 FOR BLOCK SIZE OF 64×64. THE DETECTION ACCURACIES A/B/C IN THE TABLE ARE OBTAINED BY THE BACM-BASED (A), ICA-BASED (B) AND PROPOSED (C) METHOD RESPECTIVELY.THE VALUE IN BOLD DENOTES THE BEST DETECTION ACCURACY AMONG THE THREE METHODS.

QF_1\QF_2	50	60	70	80	90
50	52.8/50.6/**56.7**	51.5/51.4/**59.0**	54.4/51.6/**67.1**	56.3/54.1/**76.1**	62.2/60.3/**88.4**
60	51.3/49.9/**54.4**	50.5/50.2/**57.7**	51.4/50.3/**60.6**	53.9/53.6/**70.5**	59.0/58.6/**85.7**
70	51.1/50.0/**53.7**	50.3/50.1/**54.5**	51.3/51.1/**58.2**	52.7/52.3/**64.4**	56.3/56.5/**80.1**
80	50.4/50.6/**52.3**	49.6/50.5/**53.4**	51.6/50.2/**54.7**	50.4/51.5/**57.4**	54.3/53.8/**72.0**
90	50.4/49.6/**50.9**	49.8/49.8/**50.5**	49.8/48.5/**52.0**	49.3/50.9/**54.4**	50.2/50.3/**58.3**

TABLE II DETECTION ACCURACY (%) WITH VARIOUS VALUES OF QF_1 AND QF_2 FOR BLOCK SIZE OF128×128. THE DETECTION ACCURACIES A/B/C IN THE TABLE ARE OBTAINED BY THE BACM-BASED (A), ICA-BASED (B) AND PROPOSED (C) METHOD RESPECTIVELY.THE VALUE IN BOLD DENOTES THE BEST DETECTION ACCURACY AMONG THE THREE METHODS.

QF_1\QF_2	50	60	70	80	90
50	51.8/52.8/**63.3**	52.6/55.3/**68.5**	58.2/60.1/**76.3**	64.3/67.2/**85.7**	74.3/76.9/**95.0**
60	50.9/51.3/**61.7**	51.3/52.4/**65.2**	54.1/56.9/**70.2**	58.5/62.8/**80.9**	69.5/73.8/**93.7**
70	50.0/50.8/**58.1**	51.2/50.8/**61.6**	52.6/53.7/**64.5**	53.8/58.7/**75.3**	62.8/71.1/**89.6**
80	50.1/49.9/**54.5**	51.4/50.6/**57.3**	50.1/50.0/**60.3**	51.0/52.6/**65.2**	57.3/64.9/**81.4**
90	50.2/50.1/**51.3**	50.2/49.8/**52.6**	49.4/49.0/**54.6**	49.9/51.3/**60.8**	52.7/51.0/**66.1**

TABLE III DETECTION ACCURACY (%) WITH VARIOUS VALUES OF QF_1 AND QF_2 FOR BLOCK SIZE OF 256×256. THE DETECTION ACCURACIES A/B/C IN THE TABLE ARE OBTAINED BY THE BACM-BASED (A), ICA-BASED (B) AND PROPOSED (C) METHOD RESPECTIVELY.THE VALUE IN BOLD DENOTES THE BEST DETECTION ACCURACY AMONG THE THREE METHODS.

QF_1\QF_2	50	60	70	80	90
50	53.2/58.8/**75.6**	56.4/64.6/**80.0**	65.6/71.2/**86.1**	75.8/80.2/**95.0**	84.4/88.2/**99.0**
60	50.9/54.9/**70.2**	55.0/61.0/**76.1**	58.2/67.0/**80.3**	67.8/77.7/**91.0**	80.6/87.8/**98.2**
70	50.8/51.0/**65.4**	50.7/54.6/**69.0**	52.1/56.9/**75.7**	59.0/69.9/**84.0**	74.2/85.3/**96.2**
80	49.6/51.2/**58.9**	49.4/50.3/**63.0**	50.0/49.9/**67.3**	53.6/55.4/**74.6**	64.9/76.3/**92.0**
90	49.9/50.9/**53.5**	50.7/49.5/**55.9**	50.0/49.6/**58.6**	50.0/52.6/**65.3**	54.4/51.4/**73.9**

From Table I, II and III, the observations can be made as follows:

(1) Due to the three kinds of errors (i.e., quantization error, truncation error and rounding error) in JPEG compression and decompression procedure, the artifacts caused by the SD-JPEG compression are extensive. Difference JPEG 2-D arrays along four directions are used to enhance them. This is one reason why the proposed method can achieve good performance.

(2) The larger the QF_2-QF_1 is, the higher the detection accuracy is. The reason for this is that the lower quality factor of the first JPEG compression will leave more traces of the compression history and the higher quality factor of the second JPEG compression will introduce less distortion to the image, which makes the traces left by the first compression easier to be detected.

(3) For the same QF_2-QF_1, the detection accuracy increase dramatically with the increase of block size. It is because that all the approaches are based on different statistical models, larger image block results in better statistical performance.

(4) Among all the three methods, the proposed method indeed outperforms the others especially when the image block size is small. The reasons are that the BACM-based method focuses on histogram in the spatial domain, which is the first order statistics. The ICA-based method measures the independency of BDCT coefficients by the objective function according to 64 different segmentation schemes. Note that the BACM-based and ICA-based methods will lead to the insufficient statistics when the image block size is small. The proposed method applies transition probability matrix to model the thresholded difference JPEG 2-D arrays, which is a kind of the second order statistics, so more information conducted by SD-JPEG compression can be caught by the proposed method than the BACM-based and ICA-based methods when the image block size is small.

(5) The detection performance of the proposed method can be improved when QF_1 is larger than QF_2 with the increase of block size. The reason for this is that second order statistics of the global JPEG 2-D array lead to better statistical performance for larger block size. However, the detection performance of the BACM-based and ICA-based method is always close to the performance of random guessing (50%) when $QF_1 > QF_2$.

B. *Detection of Copy-paste Image Forgery*

Two examples of JPEG image forgery detection using the proposed method are shown in Fig.4 and 5. All the original images are come from the NRCS image database. The composite JPEG images are generated as described in Section II. Note that the quality factor QF_1 is unknown in most cases, but the quality factor QF_2 can be obtained easily via comparison of the current quantization table. Such problem can be solved by training an appropriate SVM classifier. In this experiment, $\{QF_1, QF_2\} = \{50, 90\}$ and $\{60, 90\}$ are investigated for Fig.4 and 5 respectively. In order to detect the following composite images, the quality factor QF_1 can be randomly chosen in the intervals [50, 90], QF_2 is equal to 90. 1000 pairs of image blocks with the size 128×128 (i.e. 1000 SD-JPEG compressed blocks and 1000 single JPEG compressed blocks) are generated from the UCID database. SD-JPEG compressed blocks are generated with the quality factor QF_1, QF_2 and random coordinate shifts. Single JPEG compressed blocks are generated only with the quality factor QF_2 and same coordinate shifts. Each pair of SD-JPEG compressed block and single JPEG compressed block have the same size and image contents for each bmp images. Half of them, i.e. 500 SD-JPEG compressed blocks and 500 single JPEG compressed blocks are randomly chosen to train the SVM classifier and others for testing. The classifier is selected to yield a 1% false positive rate (a single JPEG compressed block is incorrectly classified as a SD-JPEG compressed block). The trained classifier is applied to identify all the blocks of the following composite images. A median filter is used to further remove the isolated blocks which are falsely identified as the SD-JPEG compressed blocks. The detected tampered regions are masked in black as shown in Fig.4 (c) and 5 (c).

From Fig.4 and 5, we notice that the proposed method can detect and locate the tampered regions accurately. It should be pointed out that the proposed approach can detect and locate the tempered regions with the assumption that the first JPEG compression uses a lower quality factor than the second JPEG compression. It will fail when the first JPEG compression uses a higher quality factor than the second compression.

(a) (b) (c)

Fig.4. An example of JPEG image forgery detection: (a) composite JPEG image; (b) ground truth of (a), the tampered region are bounded by the red contour;(c) detected tampered region by the proposed method.

(a) (b) (c)

Fig.5. Another example of JPEG image forgery detection: (a) composite JPEG image;(b) ground truth of (a), the tampered region are bounded by the red contour;(c) detected tampered region by the proposed method.

V. CONCLUSIONS

In this paper, an effective SD-JPEG compression detection method based on Markovian transition probability matrix is proposed. Elements in a JPEG 2-D array are often correlated for a single JPEG compressed block. The SD-JPEG compression artifacts disturb the JPEG 2-D array, weakening the correlation among the elements of the JPEG 2-D array. Such artifacts can be enhanced by the difference JPEG 2-D arrays along four directions (i.e., horizontal, vertical, main diagonal and minor diagonal). The transition probability matrix, which characterizes the Markov random process, is used to model these thresholded difference JPEG 2-D arrays. All the elements of these transition probability matrices are served as features to train the SVM classifier to discriminate the SD-JPEG compression from the single JPEG compression. The experimental results have proved that the proposed method is more effective than the BACM-based and ICA-based methods especially when the image block size is small. We also performed experiments to detect and locate the tampered regions of two composite JPEG images. The results show that the proposed approach can detect and locate the tempered blocks accurately when the first JPEG compression uses a lower quality factor than the second JPEG compression. The future work is to focus on the theoretical analysis of the SD-JPEG compression, improving detection accuracy to make the proposed approach closer to practical application, and exploring new SD-JPEG compression detection schemes.

ACKNOWLEDGEMENTS

The work described in this paper is funded by National Science Foundation of China (61071152, 60702043), 973 Program (2010CB731403, 2010CB731406) of China, Shanghai Educational Development Foundation and Project of Beijing Key Laboratory of Communication and Information System of Beijing Jiao Tong University.

REFERENCES

[1] M. Barni, “Effectiveness of exhaustive search and template matching against watermark desynchronization,” *IEEE Signal Processing Letters*, vol.12, no. 2, pp. 158–161, February 2005.

[2] M. L. Miller, G. J. Doerr, and I. J. Cox, “Applying informed coding and embedding to design a robust high-capacity watermark,” *IEEE Trans. on Image Processing*, vol. 13, no. 6, pp. 792–807, June 2004.

[3] A. C. Popescu and H. Farid, “Exposing digital forgeries by detecting traces of re-sampling”, *IEEE Trans. on Signal Processing*, vol. 53, no.2, pp.758-767, February 2005.

[4] H. Farid, “A survey of image forgery detection,” *IEEE Signal Processing Magazine*, vol. 2, no. 26, pp. 16–25, March 2009.

[5] X. Zhao, J. Li, S. Li and S. Wang, “Detecting digital image splicing in chroma spaces,” *Proc. of International Workshop on Digital Watermarking (IWDW 2010)*, Seoul, Korea, pp.12-22, October 2010.

[6] J. Lukas and J. Fridrich, “Estimation of primary quantization matrix in double compressed jpeg images,” *Proc. of Digital Forensic Research Workshop*, Cleveland, Ohio, USA, pp.5-8, August 2003.

[7] A. C. Popescu, “Statistical tools for digital image forensics,” PhD thesis, Dartmouth College, Hanover, NH, USA, December 2004. (Advised by H. Farid).

[8] D. Fu, Y. Q. Shi and Q. Su, “A generalized benford's law for jpeg coefficients and its applications in image forensics,” *Proc. of SPIE Electronic Imaging, Security and Watermarking of Multimedia Contents IX*, San Jose, CA, vol. 6505, pp. 1L1-1L11, 2007.

[9] B. Li, Y. Q. Shi and J. Huang, “Detecting doubly compressed JPEG images by using Mode Based First Digit Features,” *Proc. of IEEE Workshop on Multimedia Signal Processing (MMSP2008),* Cairns, Queensland, Australia, pp.730-735, October 2008.

[10] C. Chen, Y. Q. Shi and W. Su, “A machine learning based scheme for double JPEG compression detection,” *Proc. of IEEE International Conference on Pattern Recognition (ICPR 2008)*, Tampa, Florida, USA, pp.1-4, December 2008.

[11] W. Luo, Z. Qu, J. Huang and G. Qiu, “A novel method for detecting cropped and recompressed image block,” *Proc. of IEEE International Conference on Acoustic, Speech and Signal Processing (ICASSP' 07)*, Honolulu, Hawaii, vol.2, pp. 217-220, April 2007.

[12] Y. L. Chen and C. T. Hsu, “Image tampering detection by blocking periodicity analysis in JPEG compressed images,” *Proc. of IEEE International Workshop on Multimedia Signal Processing (MMSP 2008)*, Cairns, QLD, Australia, pp. 803-808, November 2008.

[13] Z. Qu, W. Luo and J. Huang, “A convolutive mixing model for shifted double JPEG compression with application to passive image authentication,” *Proc. of IEEE International Conference on Acoustic, Speech and Signal Processing*, Las Vegas, Nevada, USA, pp. 1661-1664, March 2008.

[14] P. Sallee, Matlab JPEG Toolbox [Online]: http://www.philsallee.com/jpegtbx/index.html.

[15] G. Schaefer and M. Stich, “UCID: An uncompressed color image database,” *Proc. SPIE Storage and Retrieval Methods and Applications for Multimedia*, San Jose, CA, vol. 5307, pp. 472–480, 2003.

[16] NRCS Photo Gallery 2005 [Online]: http://photogallery.nrcs. usda.gov/.

[17] Chang. C. C. and Lin, C. J. LIBSVM: A Library for Support Vector Machines. 2001. http://www.csie.ntu.edu.tw/cjlin/libsvm.