Detection of Misaligned Cropping and Recompression with the Same Quantization Matrix and Relevant Forgery

Qingzhong Liu
Department of Computer Science
Sam Houston State University
Huntsville, TX 77341, USA
E-mail: liu@shsu.edu

ABSTRACT
Image tampering, being widely facilitated and proliferated by today's digital techniques, is increasingly causing problems concerning the authenticity of digital images. As one of the most favorable compressed media, JPEG image can be easily tampered without leaving any visible clues. JPEG-based forensics, including the detection of double compression, interpolation, rotation, etc, has been actively performed. However, the detection of misaligned cropping and recompression, with the same quantization matrix that was once used to encode original JPEG images, has not been effectively expressed or ignored to some extent. Aiming to detect such manipulations for forensics purpose, in this paper, we propose an approach based on block artifacts caused by the manipulation with JPEG compression. Specifically, we propose a shift-recompression based detection method to identify the inconsistency of the block artifacts in doctored JPEG images. The learning classifiers are applied for classification. Experimental results show that our approach is very promising to detect misaligned cropping and recompression with the same quantization matrix and greatly improves the existing methods. Our detection method is also very effective to detect relevant copy-paste and composite forgery in JPEG images.

Categories and Subject Descriptors
I 4.9 **[Image Processing and Computer Vision**]: Applications; K.6.m **[Miscellaneous**]: Insurance and Security.

General Terms
Algorithms and Security

Keywords
Forgery, misaligned cropping, quantization matrix, shift-recompression, block artifacts, SVM, LibSVM, logistic regression, tampering, image forensics, shift-recompression based reshuffle characteristic

1. INTRODUCTION
While being widely adopted, transmitted, and enjoyed, digital multimedia can be easily manipulated without leaving an obvious clue. In recent years, multimedia forensics has emerged as a new discipline as it has important applications in protecting public safety and national security, as well as impacts to our daily life. In multimedia forensics, steganalysis and forgery detection are two interesting areas with broad impact to each other. While multiple promising and well-designed steganalysis methods have been proposed and several steganographic systems have been successfully steg-analyzed [12, 16, 18-20, 23, 24, 33], it seems that the advance in forgery detection falls behind.

Today's digital techniques make it easy to widely spread digital multimedia, wherein JPEG image is one of the most popular digital images in our daily life. While we enjoy huge volumes of JPEG images in digital format, our traditional confidence in the integrity via our eyes and ears has also been undermined since doctored pictures, video clips, and audio streams are easily manipulated. For example, a recent state-run newspaper in Egypt published a doctored picture, attempting to create the illusion that its country's president was leading the group in Middle East peace talks in Washington DC [1].

Generally, tampering manipulation in digital media involves several basic operations, such as image resize, rotation, splicing, double compression. The detection of these fundamental manipulations and relevant forgery has been well studied [2-11, 14, 17, 21, 22, 26-32], for instance, double JPEG compression is one of most adopted manipulations. While we decode the bit stream of a JPEG image and implement the manipulation in spatial domain, and then compress the modified image back to JPEG format, if the newly adopted quantization matrix is different from the one used by original JPEG image, we say the modified JPEG image has undergone a double JPEG compression. Although JPEG based double compression does not by itself prove malicious or unlawful tampering, it is an evidence of image manipulation. The detection of double JPEG compression has been well studied [4, 22, 28, 29]. However, if a forgery is made from the image sources encoded at the same compression quality, such detection is not effective. Although Huang et al. presented a method to detect double JPEG compression with the same quantization matrix, but it cannot tell us the double-compressed JPEG image is composited or not [14].

Recently, Luo et al. designed a set of block artifact characteristics matrix features (BACM) to detect the JPEG images once cropped and recompressed [26]. Chen and Hsu analyzed the periodicity of compression artifacts for tampering detection [5]. Both methods are impressive for the detection of cropping and recompression with different quantization matrices, unfortunately, they are not effective in the detection of the cropping and recompression with the same quantization matrix, shown by the results in the reference [5]. To our knowledge, existing methods do not work well to detect doctored images with the recompression by using the same quantization matrix, which was once used to encode the image sources. Our study aims to solve this problem in the paper.

MiFor'11, November 29, 2011, Scottsdale, Arizona, USA.

In section 2, we propose a novel approach to detect the manipulation of misaligned cropping and recompression with the same quantization matrix for JPEG-based forensics. Section 3 presents several types of substantial experiments, including detection of cropping manipulation, copy-paste detection, and composite detection. Section 4 presents our conclusion, followed by acknowledgments in Section 5.

2. SHIFT-RECOMPRESSION-BASED DETECTION APPROACH

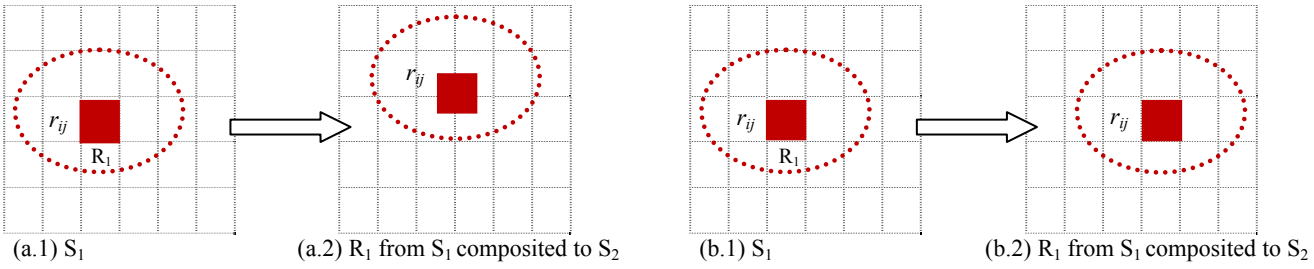
2.1 Misaligned Cropping and Recompression

To prevent a forgery manipulation on JPEG images from being detected, a crafty forgery maker may try to avoid double JPEG compression during the manipulation, since the detection of JPEG double compression has been well studied with satisfactory results. It is not difficult for a forgery maker to obtain the two source JPEG images with the same compression quality, that is, the encoding to JPEG format takes the same quantization matrix (the quantized DCT coefficients are obtained by dividing the pre-quantized DCT coefficients by the same quantization matrix table). In tampering, source JPEG images will be decoded or uncompressed to spatial domain first, and manipulation takes place in spatial domain. The doctored image will be compressed to JPEG format at the same quality, or quantizing the pre-quantized DCT coefficients by using the same quantization matrix that was once used by the source images.

We describe the manipulation operations as follows: the source images S_1 and S_2, with the DCT coefficients quantized by using quantization table QT. Te create a forgery from S_1 and S_2, both source images are uncompressed and shown in spatial domain, a region of interest R_1 from S_1 is copied and pasted to S_2. Modified S_2 is compressed to JPEG format, the DCT coefficients are quantized by using the same quantization table QT.

As shown by Figure 1, the original region R_1 consists of several 8×8 JPEG-compression blocks r_{ij} in S_1. Assuming region R_1 randomly pasted in S_2, then original JPEG-compression blocks r_{ij} will be reshuffled with the neighboring 8×8 blocks (misaligned cropping and recompression) at a high probability (63/64 = 98.4%) as new 8×8 blocks,; if the block r_{ij} will not be reshuffled with the neighboring 8×8 blocks, it will be recompressed by itself as an entire 8×8 block (aligned cropping and recompression), such manipulation hits the probability of 1/64.

If S_1 and S_2 are encoded by using different quantization matrices, or if S_1 and S_2 are encoded with the same quantization matrix but the recompression shown in Figure 1 uses different quantization matrix, then double compression takes place. It is easy for us to detect such double compression. However, if S_1, S_2, and composited image are all encoded with the same quantization matrix, the detection of such compositing manipulation has not been well dealt with and ignored to some extent so far.



(a.1) S_1 (a.2) R_1 from S_1 composited to S_2 (b.1) S_1 (b.2) R_1 from S_1 composited to S_2

(a) misaligned cropping and recompression (b) aligned cropping and recompression

Figure 1. Two types of JPEG-based composition. Misaligned cropping and recompression (a) occurs at a high probability.

2.2 Shift-Recompression-Based Approach

Inspired by the method to detect double JPEG compression [10] and the methodology of self-calibration that was used in steganalysis [12, 25], we propose a shift-recompression-based algorithm to detect misaligned cropping and recompression with the same quantization matrix in JPEG images.

We surmise that the reshuffle, shown in Figure 1(a), will leave clues for us to detect such manipulation behaviors in the final doctored JPEG image, although double JPEG compression has been avoided. Accordingly we design a shift-recompression based algorithm to identify the inconsistency of block artifacts due to the reshuffle behaviors, and finally identify the forged area in encoded JPEG formats. The algorithm is described as follows:

SRSC feature extraction algorithm

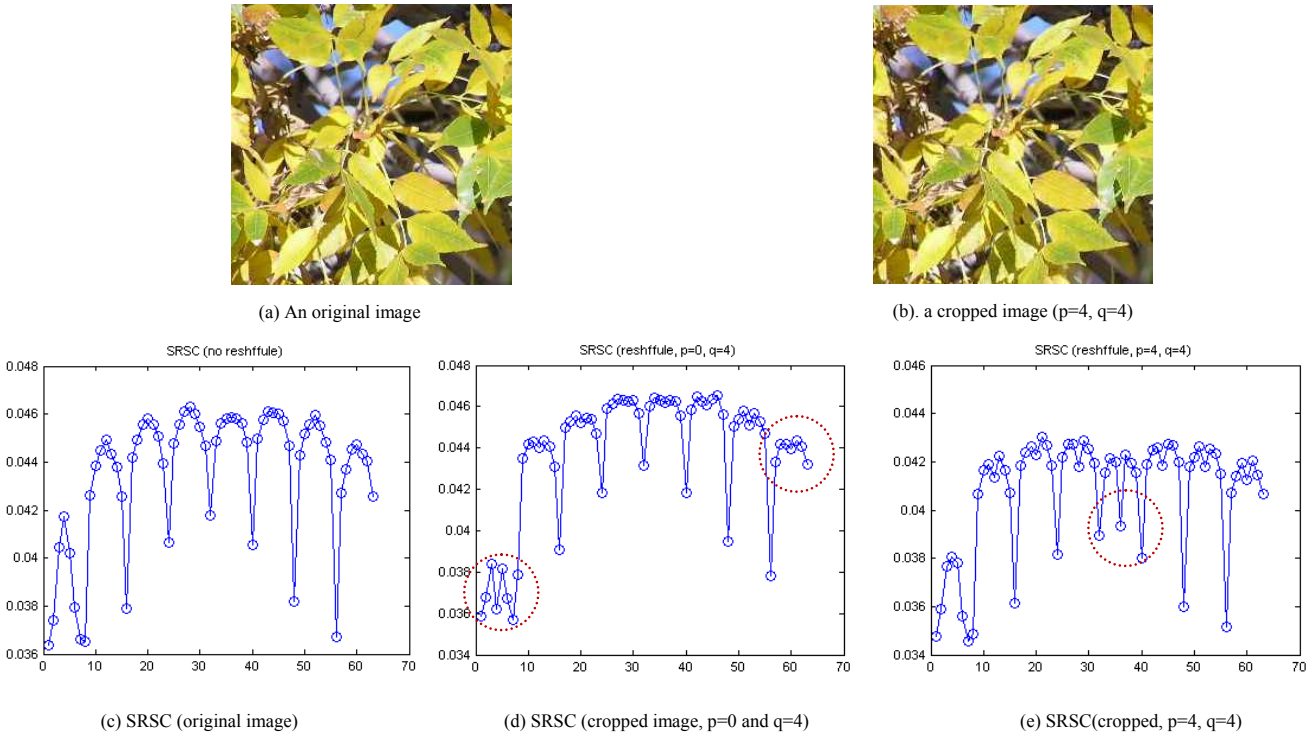
1. Decode an JPEG image under examination to spatial domain, which is denoted by matrix S(i,j) (i=1, 2, …, M; j = 1, 2 …, N);
2. Shift the matrix S(i,j) by d_1 rows and d_2 columns in the spatial domain, $(d_1, d_2) \in \{(0,1), \ldots, (0,7), (1,0), \ldots, (7,7)\}$ and generate a shifted spatial image S′(d_1, d_2), S′(d_1, d_2) = S(i- d_1, j- d_2)(i= d_1+1, d_1+2, …, M; j= d_2+1, q+ d_2, …, N);
3. Compress the shifted spatial image S′(d_1, d_2) to JPEG format at the same quality factor;
4. Decode the shifted JPEG image to spatial domain, denoted by a matrix S′′(d_1, d_2);
5. Calculate the difference D (d_1, d_2)= S′(d_1, d_2) - S′′(d_1, d_2);
6. **S**hift-recompression based **Re**S**h**uffle **C**haracteristic features (**SRSC**) on the region of interest R, **SRSC**$_R$ are defined by:

$$SRSC_R(d_1,d_2) = \frac{\sum |D_R(d_1,d_2)|}{\sum |S'_R(d_1,d_2)|}, \quad (1)$$

Where $(d_1,d_2) \in \{(0,1),\ldots,(0,7),(1,0),\ldots,(7,7)\}$, total 63 features, for each R.

If an image was cropped with the misalignment by p rows and q columns, mod (p, 8) ≠ 0 or mod (q, 8) ≠0, 0≤p≤8, 0≤q≤8, and then recompressed at the same quality level to the original JPEG image, we expect that the SRSC features will be distinct due to the misalignment, and the values of p and q can be determined by the SRSC features. The example shown in Figure 2 confirms our

conjecture and preliminarily validates our algorithm. Figure 2 shows an original JPEG image (a) and a cropped image with misalignment p=4 and q=4 (b). The SRSC features from original image and two cropped are shown in (c), (d), and (e). The circles highlight part differences of SRSC features compared to the SRSC features extracted from the JPEG image without cropping.



(a) An original image

(b). a cropped image (p=4, q=4)

(c) SRSC (original image)

(d) SRSC (cropped image, p=0 and q=4)

(e) SRSC(cropped, p=4, q=4)

Figure 2. A comparison of SRSC features from an original JPEG image and three cropped JPEG images. X-label shows the SRSC feature index and y-label indicates the value.

3. EXPERIMENTS

3.1 Detection of Misaligned Cropping and Recompression—Binary Classification

To test our proposed shift-recompression based SRSC features, we select 5150 singly compressed JPEG images at the quality factor 85, and 5150 singly compressed JPEG images at quality factor 40. Respectively, we cropped these JPEG images by all misalignment combinations (p, q), from (0,1), …, (0,7); (1,0), …., (1,7); …. to (7,7), total 63 combinations, and produced 5150×63 =324450 cropped JPEG images at the quality 85 from the same quality JPEG images, and 324450 cropped JPEG image at quality 40 from the same quality JPEG images. In a fair manner, since these cropped JPEG images are undergone twice JPEG compression with the same quantization matrix, the singly compressed JPEG images are also uncompressed and then recompressed by using the same quantization matrix. Then we extract SRSC features from all these JPEG images. Support vector machines [34] are employed in our detection to discriminate each type of misaligned cropping from no cropping, which is a binary classification from the perspective of pattern recognition. In our experiments, we apply two popular SVM technique, SVMlight [15], and LibSVM [35], with linear kernel, polynomial kernel, and RBF kernel individually to the features, for training and testing. The ratio of training to testing is 50% to 50%, fifty experiments are operated in each type of detection in these binary classifications. In each experiment, training feature set is randomly selected and remaining feature set is selected for validation. Testing results can be divided into true positive (TP), false negative (FN), false positive (FP), and true negative (TN), based on the ground truth and prediction results. In our experiments, the classification results by using LibSVM are generally better than SVMlight. And hence, we only list the results using LibSVM in Table 1. We calculate testing accuracy by a half of the sum of true positive rate and true negative rate, or 0.5*(TP/(TP+FN)+TN/(TN+FP)).

A set of block artifact characteristics matrix features (BACM) has been proposed to detect the JPEG images once cropped and recompressed [26], a recent periodicity analysis of compression artifacts for tampering detection takes advantage of BACM and the detection results on misaligned cropping and recompression with the same quantization matrix is not effective, shown by the results in [5], and the feature extraction algorithm has not been clearly illustrated in [5], and hence in our comparative study, we only compare our approach to BACM feature set.

The experimental results shown by Tables 1-1 and 1-2 clearly demonstrate the incomparable superiority of our approach. The BACM-based detection performance is not well and the results are not reliable, but SRSC-based approach is very impressive, especially with the use of LibSVM, most average testing accuracy values are over 98%, while the detection accuracy values based on BACM sway around 60%.

Table 1-1 Binary classification accuracy on testing sets by using LibSVM with linear, polynomial and RBF kernels, the best average testing accuracy value by using these kernels is shown on the following Table (Q=40).

p \ q / feature set		0	1	2	3	4	5	6	7
0	BACM		64.5%	55.7	54.8	54.2	55.3	56.9	62.2
	SRSC		**96.6**	**99.0**	**98.9**	**99.3**	**98.8**	**98.8**	**96.9**
1	BACM	61.9	57.5	55.7	55.9	56.5	53.9	54.9	57.2
	SRSC	**95.7**	**96.7**	**99.1**	**99.4**	**99.4**	**99.3**	**99.0**	**97.5**
2	BACM	56.9	57.9	57.5	56.3	57.7	55.3	56.2	57.2
	SRSC	**98.6**	**99.0**	**99.5**	**99.5**	**99.5**	**99.6**	**99.5**	**99.0**
3	BACM	55.7	58.0	58.2	59.5	56.9	53.9	56.8	56.9
	SRSC	**98.5**	**99.4**	**99.4**	**98.8**	**98.9**	**98.7**	**99.5**	**99.3**
4	BACM	54.3	57.6	58.0	58.1	57.1	54.1	55.1	56.2
	SRSC	**98.9**	**99.5**	**99.5**	**98.8**	**98.9**	**98.6**	**99.4**	**99.4**
5	BACM	59.9	56.5	58.6	58.6	55.8	53.9	56.8	57.0
	SRSC	**98.7**	**99.4**	**99.5**	**98.9**	**98.9**	**98.7**	**99.3**	**99.2**
6	BACM	60.0	56.8	57.0	55.2	55.6	54.1	55.6	56.2
	SRSC	**98.7**	**99.0**	**99.4**	**99.5**	**99.5**	**99.4**	**99.3**	**98.9**
7	BACM	60.9	57.9	56.3	56.8	55.9	56.6	54.7	56.3
	SRSC	**95.9**	**97.3**	**99.1**	**99.4**	**99.4**	**99.1**	**99.1**	**97.5**

Table 1-2 Binary classification accuracy on testing sets by using LibSVM with linear, polynomial and RBF kernels, the best average testing accuracy value by using these kernels is shown on the following Table (Q=85).

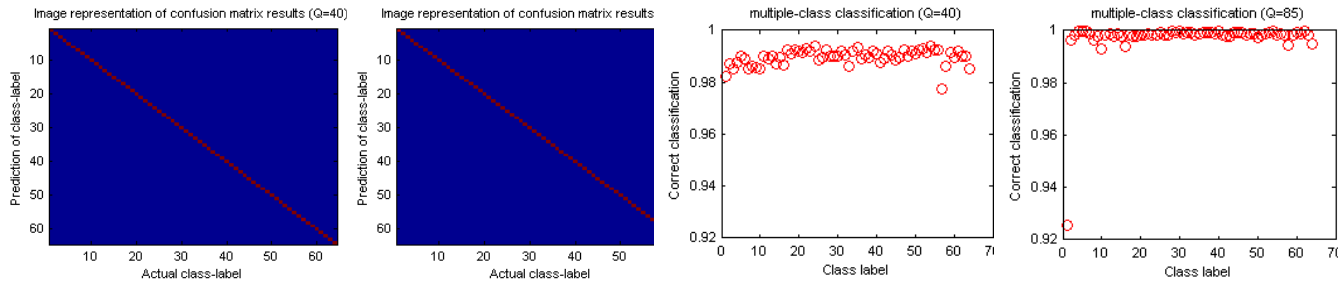
p \ q / feature set		0	1	2	3	4	5	6	7
0	BACM		63.6%	62.6	62.2	61.4	65.2	65.2	62.5
	SRSC		**99.1**	**99.3**	**99.6**	**99.7**	**99.4**	**99.1**	**98.8**
1	BACM	62.5	58.0	59.2	59.8	61.2	60.9	59.6	58.8
	SRSC	**98.6**	**98.7**	**99.0**	**99.2**	**99.3**	**99.1**	**98.8**	**98.6**
2	BACM	60.4	58.5	58.1	59.1	59.6	60.5	59.3	58.5
	SRSC	**99.0**	**99.1**	**98.9**	**99.1**	**99.2**	**98.9**	**98.6**	**98.9**
3	BACM	62.3	59.3	58.9	60.5	60.6	63.7	60.2	59.9
	SRSC	**99.2**	**99.3**	**99.1**	**98.9**	**98.9**	**98.6**	**98.7**	**99.2**
4	BACM	60.6	61.0	61.0	63.2	64.7	65.4	62.3	61.2
	SRSC	**99.2**	**99.3**	**99.1**	**98.8**	**98.9**	**98.6**	**98.8**	**99.2**
5	BACM	65.9	62.5	62.8	63.3	66.4	66.0	63.7	63.5
	SRSC	**99.1**	**99.1**	**98.9**	**98.7**	**98.7**	**98.5**	**98.7**	**98.9**
6	BACM	63.1	58.8	58.7	60.5	63.4	63.7	60.5	59.4
	SRSC	**98.9**	**98.9**	**98.7**	**98.9**	**99.0**	**98.7**	**98.3**	**98.6**
7	BACM	62.9	59.4	59.1	59.5	61.6	62.3	59.7	58.8
	SRSC	**98.5**	**98.8**	**99.1**	**99.1**	**99.2**	**99.0**	**98.6**	**98.6**

3.2 Detection of Misaligned Cropping and Recompression—Multiple-Class Classification

In Figure 1 (a), if S_2 is cropped, for example, the pixels on the boundary are stripped off, then the region R_1 from S_1 is composited to S_2, and the doctored image is compressed with the same quantization matrix, in this case, how do we identify the forged area in the compositing? The binary classification shown in section 3.1 is not good enough. If we can identify the misalignment of S_2 from the misalignment of R_1, then we can reveal the different cropping manipulations and locate the forged area in the compositing.

In this type of experiments, we select 2000 singly compressed JPEG images at the quality factor 85, and 2000 singly compressed JPEG images at quality factor 40. Respectively, we cropped these JPEG images with the all displacement combinations (p, q), from (0,1), …, (0,7); (1,0), …., (1,7); …. to (7,7), total 63 combinations, and produced 2000×63 =126000 cropped JPEG images at the quality 85, and 126000 cropped JPEG image at quality 40. In this type experiment, a logistic regression classifier [13] is employed to the features, corresponding to quality factor 85, and 40, respectively, for training and testing. The ratio of training to testing is 50%: 50% and 100 experiments are operated at each quality factor for multiple-class classification, or identification of the cropping and the misalignment distances. In each experiment, training feature set is randomly selected and remaining feature set is for validation.

We obtained the confusion matrix of average accuracy over the 100 experiments in the multiple-class classification (total 64 labels, containing 64×64 = 4,096 average accuracy values at each quality factor), the accuracy values are shown in image format by Figure 3 (a) and (b). The average accuracy values along the diagonal direction or correct recognition for each cropping type of JPEG images are given by Figure 3 (c) and (d). The x-label indicates the class label (class 1 represents original image, class 2 to class 64 denote the misalignment distance coordinates from (0,1) to (7,7), respectively), and y-label shows the correct classification for each class in all combinations.



(a) (b) (c) (d)

Figure 3. Confusion matrix of average accuracy values on the testing (first row) and correct classification for each type of displaced JPEG images (second row) in multiple-class classification by using logistic regression classifier

3.3 Shift-Recompression Based Detection of Relevant Copy-Paste and Composite Forgery

To create copy-paste forgery and composite forgery database, we select 2000 singly compressed JPEG images at the quality factor 85 and create 2000 copy-paste JPEG forgery at central 64×64 region, and 2000 composite JPEG forgery at central 64×64 region, with a random selection of displacement and the manipulated images are encoded in JPEG format at the same compression quality to the pre-manipulated image (or by using the same quantization table). To detect such copy-paste and composite manipulations, we extract the SRSC features from different regions of an image by using the following procedure:

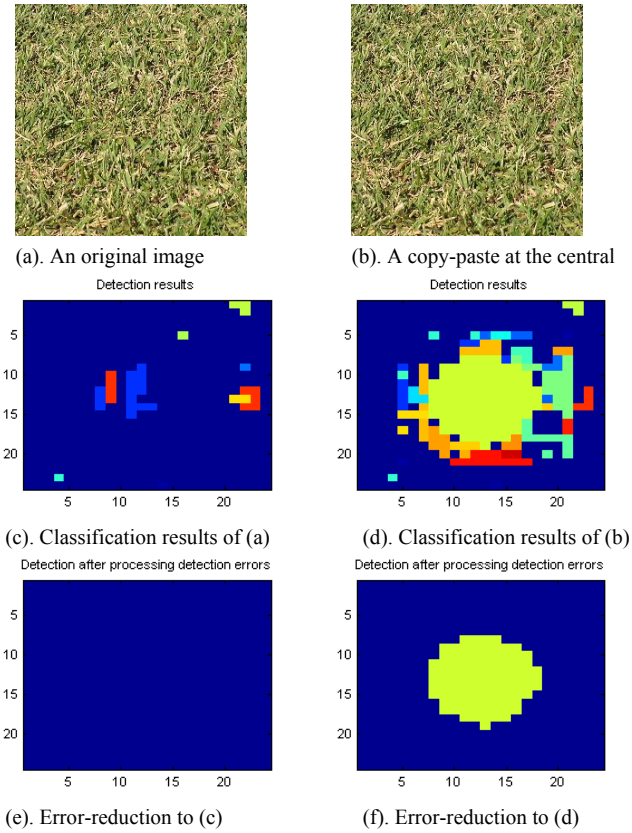
1. Extract $\mathbf{SRSC}_R$ features from each region of interest R. Let $\mathbf{R}(r_1, r_2)$ stand for the (r_1, r_2) sub-region of S′, and four horizontal and vertical neighbor regions are denoted by $\mathbf{R}(r_1\text{-}1, r_2)$, $\mathbf{R}(r_1\text{+}1, r_2)$, $\mathbf{R}(r_1, r_2\text{-}1)$, and $\mathbf{R}(r_1, r_2\text{+}1)$. We slide a window over the image under examination from the upper-left to the right-bottom, in the horizontal direction first and then in the vertical direction. Each movement of the window shifts 8 pixels. In our experiment, the window size is set 64×64, $\mathbf{R}(r_1\text{-}1, r_2)$, $\mathbf{R}(r_1\text{+}1, r_2)$, $\mathbf{R}(r_1, r_2\text{-}1)$, and $\mathbf{R}(r_1, r_2\text{+}1)$ have 87.5% overlap with $\mathbf{R}(r_1, r_2)$.
2. The multiple-class classification models are loaded to classify the features from each region, and all prediction results are organized as a two-dimensional array, in terms of the region indices.
3. Based on the class-label occurrence, we can apply another learning classifier to automatic recognition of the forged image, and the approximate forgery region will be located. The sparsely distributed class labels with the label value larger than one are very probably the classification errors (due to the high portion overlapping of the neighboring regions), therefore these areas should not be recognized as forgery. This processing may be called error-reduction or noise-removal.

Table 2 shows the detection results without error reduction process when distinguishing copy-paste and composite forgery from untouched JPEG images by using a linear LibSVM and logistic regression classifiers, respectively.

Table 2. Average detection performance over 50 times using a linear LibSVM and logistic regression

Classifier	True Negative Rate	True Positive Rate
LogitReg	99.6%	99.5%
LibSVM	99.5%	99.4%

To obtain the results in Table 2, we first predict the class-label of each sub-region of the image under examination by loading the multiple-classification models established by applying logistic regression classifier to SRSC features, described in Section 3.2, the number of total possible class labels is 64. The occurrence probability of each class label from the prediction forms the input of feature vector. In each experiment, we randomly select 60% feature sets for training and other 40% feature sets are tested. Fifty experiments are performed for each testing.



(a). An original image (b). A copy-paste at the central
(c). Classification results of (a) (d). Classification results of (b)
(e). Error-reduction to (c) (f). Error-reduction to (d)

Figure 4. An illustration of forgery detection using SRSC features and logistic regression classifier.

The classification accuracy shown in Table 2 does not employ error-reduction process. Figure 4 shows an example with error-reduction to detect two JPEG images with an untouched JPEG image on the left (a) and a paste forgery at the central of the image on the right (b), with the copy source from the upper-left, at the same quantization factor. Figure 4(c) and (d) are the image representation of detection results with the sub-region indices

demonstrated by x-axis and y-axis. Figure 4(e) and (f) are the final results after error-reduction or noise-removal.

4. CONCLUSIONS

In this paper, we propose a shift-recompression-based approach to detection of misaligned cropping and recompression with the same quantization matrix and relevant forgery in JPEG images, by revealing the inconsistency of the block artifacts caused by the manipulation. The classifiers SVM and logistic regression are applied to the SRSC features for the detection. Experimental results show that our approach greatly outperforms relevant existing methods in JPEG-based cropping and recompression detection. Shift-recompression based approach is also very promising to detect relevant copy-paste and composite forgery in JPEG images.

5. ACKNOWLEDGMENTS

This project was supported by Award No. 2010-DN-BX-K223 awarded by the National Institute of Justice, Office of Justice Programs, U.S. Department of Justice. The opinions, findings, and conclusions or recommendations expressed in this publication/program/exhibition are those of the authors and do not necessarily reflect those of the Department of Justice. Part support for this study from SHSU research office is greatly appreciated. We are also grateful to Mrs. Sarla Mile for her proofreading.

6. REFERENCES

[1] http://www.npr.org/blogs/thetwo-way/2010/09/17/129938169/doctored-photograph-hosni-mubarak-al-ahram-white-house-obama-mideast-peace-talks

[2] Bayram S, Sencar HT and Memon N (2009). An efficient and robust method for detecting copy-move forgery. *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing*, April 19-24, 2009.

[3] Chen W, Shi YQ and Su W (2007). Image splicing detection using 2-D phase congruency and statistical moments of characteristic function. *Proc. SPIE*, Vol. 6505, 65050R (2007); DOI:10.1117/12.704321.

[4] Chen C, Shi YQ and Su W (2008). A machine learning based scheme for double JPEG compression detection. *ICPR* 2008: 1-4.

[5] Chen Y and Hsu C (2011). Detecting recompression of JPEG images via periodicity analysis of compression artifacts for tampering detection. *IEEE Transactions on Information Forensics and Security*, 6(2): 396-406.

[6] Dirik A E and Memon N (2009). Image tamper detection based on demosaicing artifacts. *Proc. IEEE ICIP'09*, November 2009.

[7] Farid H (1999). Detecting digital forgeries using bispectral analysis. AI Lab, *Massachusetts Institute of Technology, Tech. Rep. AIM-1657*, 1999.

[8] Farid H (2006). Digital image ballistics from JPEG quantization. Dept. Comput.Sci., *Dartmouth College, Tech. Rep*. TR2006-583, 2006.

[9] Farid H (2009). Image forgery detection, a survey. *IEEE Singal Processing Magazine*, March 2009, 16-25.

[10] Farid H (2009). Exposing digital forgeries from JPEG ghosts. *IEEE Transactions on Information Forensics and Security*, 4(1):154-160.

[11] Fridrich J, Soukal D and Lukás J (2003). Detection of copy move forgery in digital images. *Proc. Digital Forensic Research Workshop*, Aug. 2003.

[12] Fridrich J (2004). Feature-based steganalysis for JPEG images and its implications for future design of steganographic schemes. *Lecture Notes in Computer Science,* 3200, 67--81.

[13] Hilbe JM (2009). *Logistic Regression Models*. Chapman & Hall/CRC Press. ISBN 978-1-4200-7575-5.

[14] Huang F, Huang J and Shi Y (2010). Detecting double JPEG compression with the same quantization matrix. *IEEE Transactions on Information Forensics and Security*, 5(4):848-856.

[15] Joachims T (2000). Estimating the generalization performance of a SVM efficiently. *Proc. of the International Conference on Machine Learning,* Morgan Kaufman, 2000.

[16] Kodovsky J and Fridrich J (2009). Calibration revisited. *Proceedings of the 11th ACM Multimedia and Security Workshop*, Princeton, NJ, September 7-8, 2009.

[17] Kirchner M (2008). Fast and reliable resampling detection by spectral analysis of fixed linear predictor residue. *Proc. 10th ACM Multimedia and Security Workshop*, pp. 11–20.

[18] Liu Q, Sung AH and Qiao M (2009). Improved detection and evaluation for JPEG steganalysis. *Proc. ACM Multimedia* 2009*,* 873--876.

[19] Liu Q, Sung AH and Qiao M (2009). Novel stream mining for audio steganalysis. *Proc. 17th ACM Multimedia*, pp. 95-104, 2009.

[20] Liu Q, Sung A H and Qiao M (2009). Temporal derivative based spectrum and mel-cepstrum audio steganalysis. *IEEE Transactions on Information Forensics and Security*, 4, 3, 359--368.

[21] Liu Q and Sung AH (2009). A new approach for JPEG resize and image splicing detection. *Proc. ACM Multimedia Workshop on Multimedia in Forensics* 2009, pp. 43-47.

[22] Liu Q, Sung AH and Qiao M (2011). A method to detect JPEG-based double compression. In *Proc. of 8th International Symposium on Neural Networks*, pp 466-476.

[23] Liu Q, Sung AH and Qiao M (2011). Neighboring joint density based JPEG steganalysis. *ACM Transactions on Intelligent Systems and Technology,* 2(2), 16:1-16.

[24] Liu Q, Sung AH and Qiao M (2011). Derivative based audio steganalysis. *ACM Transactions on Multimedia Computing, Communications and Application,* 7(3), 18:1-19.

[25] Liu Q (2011). Steganalysis of DCT-Embedding-based Adaptive Steganography and YASS, *Proc. 13th ACM Workshop on Multimedia and Security*, September 28-29, 2011, Buffalo, NY.

[26] Luo W, Qu Z, Huang J and Qiu G (2007). A novel method for detecting cropped and recompressed image block. *Proc. IEEE Conf. Acoustics, Speech and Signal Processing* 2007, pp. 217–220.

[27] Pan X and Lyu S (2010). Region duplication detection using image feature matching. *IEEE Trans. on Info. Forensics and Security*, 5(4):857-867.

[28] Pevny T and Fridrich J (2008). Detection of double-compression in JPEG images for applications in steganography. *IEEE Trans. Information Forensics and Security*, 3(2):247-258, 2008.

[29] Popescu AC and Farid H (2004). Statistical tools for digital forensics. *Proc. 6th Int. Workshop on Information Hiding*, pp. 128–147.

[30] Popescu AC and Farid H (2005). Exposing digital forgeries by detecting traces of re-sampling. *IEEE Trans. Signal Processing* 53(2): 758–767.

[31] Popescu AC and Farid H (2005). Exposing digital forgeries in color filter array interpolated images. *IEEE Trans. Signal Processing* 53(10): 3948–3959.

[32] Prasad S and Ramakrishnan KR (2006). On resampling detection and its application to image tampering. *Proc. IEEE Int. Conf. Multimedia and Exposition* 2006, pp. 1325–1328.

[33] Shi YQ, Chen C and Chen W (2007). A Markov process based approach to effective attacking JPEG steganography. *Lecture Notes in Computer Science*, vol.4437, pp. 249-264.

[34] Vapnik, V. 1998. *Statistical Learning Theory*, John Wiley, 1998.

[35] http://www.csie.ntu.edu.tw/~cjlin/libsvm/