DETECTION OF NON-ALIGNED DOUBLE JPEG COMPRESSION WITH ESTIMATION OF PRIMARY COMPRESSION PARAMETERS

Tiziano Bianchi, Alessandro Piva

Università di Firenze, Dipartimento di Elettronica e Telecomunicazioni
Via S. Marta 3, 50139, Firenze - Italy

ABSTRACT

In this paper, we propose a simple yet reliable method to detect the presence of non-aligned double JPEG compression (NA-JPEG). The method is based on a single feature which depends on the integer periodicity of the DCT coefficients when the DCT is computed according to the grid of the previous JPEG compression. Even if the proposed feature is computed relying only on DC coefficient statistics, a simple threshold detector can classify NA-JPEG images with improved accuracy with respect to existing methods and on smaller image sizes. Moreover, the proposed method is able to accurately estimate the quantization step and the grid shift of the primary JPEG compression, which can be used to perform a more detailed analysis of possibly forged images.

Index Terms— Image forensics, JPEG artifacts, non-aligned double JPEG compression, threshold detector

1. INTRODUCTION

Recent technological advances have given almost everybody the possibility of recording and sharing a large amount of digital images, enabling the virtually instantaneous diffusion of visual information regarding people, events, and so on. At the same time, the large availability and the ease of use of commercial image processing tools make extremely simple to alter the content of digital images, so that the message conveyed by an image can be easily manipulated. This can be a serious problem when the content of an image is used to influence the opinion of viewers, like in tribunals, in journals, in medical applications.

Image forensics is a new discipline aiming at assessing the credibility of digital images by looking for possible inconsistencies in statistical and geometrical features, including JPEG quantization artifacts, interpolation effects, demosaicing traces [1]. An advantage of this approach is that it is not necessary to embed an explicit fingerprint in the image, making forensics tools applicable even in the case of images acquired by standard commercial cameras.

Due to the fact that a vast amount of digital images is stored in JPEG format, several forensics tools try to detect tampering by analyzing artifacts introduced by JPEG recompression. In [2], the authors propose to detect areas which have undergone a double JPEG compression by recompressing the image at different quality levels, whereas in [3] double JPEG compression is detected analyzing the statistics of DCT coefficients. The above approaches work when the DCT grids of succesive JPEG compressions are perfectly aligned. The case of non aligned double JPEG compression has been investigated in [4], by considering blocking artifacts, and in [5], where the shift of the primary JPEG compression is determined via a demixing approach.

In this paper, we propose a novel technique to detect the presence of non aligned double JPEG (NA-JPEG) compression. Differently from [4][5], in which NA-JPEG compression is detected by training a classifier on a set of features, our approach relies on a single yet powerful feature derived from the statistics of DCT coefficients, allowing us to apply a simple threshold detector. Moreover, the proposed approach is able to estimate both the quantization level and the shift of the primary compression. Such information can be used to perform a more detailed analysis of a possibly forged image.

2. FORENSIC SCENARIO AND PREVIOUS WORK

The classical scenario in which NA-JPEG compression occurs is that of image splicing. In this kind of forgery, it is assumed that a region from a JPEG image is pasted onto a host image and that the resulting image is JPEG compressed. Assuming a random placement of the forged region, there is a probability of $63/64$ that the forged region will exhibit NA-JPEG compression artifacts. This is the same scenario considered in [4][5].

An alternative scenario for NA-JPEG compression could consider the following forgery model: an original JPEG image is locally modified using a technique which disrupts JPEG compression statistics, then randomly cropped and recompressed in JPEG format. Examples of local tampering which destroys JPEG statistics could be a cut and paste from either a non compressed image or a resized image, or the insertion of computer generated content. This is similar to the sce-

nario considered in [3], with the difference that the non tampered region exhibits NA-JPEG compression instead of double JPEG compression on the same grid.

In both scenarios, a tool able to identify NA-JPEG compression could be used to discern original regions from tampered ones, for example using a segmentation of the image under test [6]. As a matter of fact, the two forensics scenarios are complementary: the only difference is in the interpretation of the NA-JPEG compressed regions as either tampered or original. We note that existing forensics tools can be used in both scenarios, so that they can be directly compared with the proposed approach.

3. PROPOSED METHOD

A JPEG decompressed image can be modeled as follows

$$\mathbf{I}_2 = \mathbf{D}_{00}^{-1}\mathcal{Q}_2(\mathbf{D}_{00}\mathbf{I}_1) + \mathbf{E}_2 = \mathbf{I}_1 + \mathbf{R}_2 \qquad (1)$$

where $\mathbf{I}_1$ is the original image, $\mathbf{D}_{00}$ models an 8×8 block DCT with the grid aligned with the upper left corner of the image, $\mathcal{Q}_2(\cdot)$ models quantization and dequantization according to the JPEG quantization table, and $\mathbf{E}_2$ is the error introduced by rounding and truncating the output values to eight bit integers. The last quantity $\mathbf{R}_2$ can be thought of as the overall error introduced by JPEG compression with respect to the original image. If we assume that the original image was previously JPEG compressed with a grid shifted by (y, x), $0 \le x \le 7$ and $0 \le y \le 7$, with respect to the upper left corner, i.e., $\mathbf{I}_1 = \mathbf{D}_{yx}^{-1}\mathcal{Q}_1(\mathbf{D}_{yx}\mathbf{I}) + \mathbf{E}_1$, then by applying a block DCT with different alignments we have

$$\mathbf{D}_{ij}\mathbf{I}_2 = \begin{cases} \mathcal{Q}_2(\mathbf{D}_{00}\mathbf{I}_1) + \mathbf{D}_{00}\mathbf{E}_2 & i = 0, j = 0 \\ \mathcal{Q}_1(\mathbf{D}_{yx}\mathbf{I}) + \mathbf{D}_{yx}\mathbf{R}_2 & i = y, j = x \\ \mathbf{D}_{ij}\mathbf{D}_{00}^{-1}\mathcal{Q}_2(\mathbf{D}_{00}\mathbf{I}_1) + \mathbf{D}_{ij}\mathbf{E}_2 & \text{elsewhere.} \end{cases}$$
$$(2)$$

When the DCT grid is aligned with the grid of either the last compression or the first compression, the DCT coefficients tend to cluster around the points of a lattice defined by the respective quantization tables. In the case of the last compression $(i = 0, j = 0)$, the spread around the lattice points is fixed and quite limited, whereas in the case of the first compression $(i = y, j = x)$ it depends on the power of $\mathbf{R}_2$, i.e., on the quality of the second JPEG compression. When the DCT grid is aligned with neither of the two compressions, DCT coefficients usually do not cluster around any lattice [7].

The main idea behind the proposed method is that of detecting NA-JPEG compression by measuring how DCT coefficients cluster around a given lattice for any possible grid shift. When NA-JPEG is detected, the parameters of the lattice also give the primary quantization table. Even if the effect described above can be measured for each DCT coefficient within an 8×8 block, in order to keep the detection simple we can consider only a subset of the DCT coefficients. In the
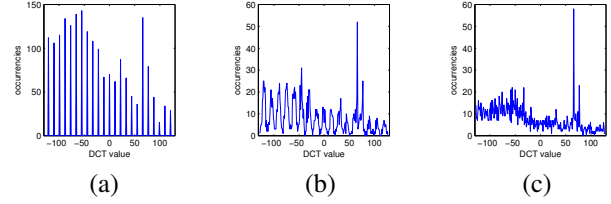
Fig. 1. Histogram of the DC coefficients for different shifts of the DCT grid: (a) $(i, j) = (0, 0)$; (b) $(i, j) = (y, x)$; (c) random shift.

following, we will consider only the DC coefficient of each block, since quantization effects will be more evident when most of the analyzed DCT coefficients are different from zero.

When a single DCT coefficient for each block is considered, clustering around a lattice can be measured by the periodicity of the histogram for an integer period. This is evident in Fig. 1, in which the histogram of the DC coefficients is shown for $(i, j) = (0, 0)$, $(i, j) = (y, x)$, and a random shift. The periodicity of the histogram can be evaluated by considering its Fourier transform at frequencies which are reciprocal of an integer value, i.e., we define

$$f_{ij}(Q) \triangleq \sum_k h_{ij}(k)e^{-j\frac{2\pi k}{Q}}, \quad Q \in \mathbb{Z}\backslash\{0\} \qquad (3)$$

where h_{ij} is the histogram of the DC coefficient for the (i, j) shift. According to (2), in the presence of NA-JPEG we will expect both $f_{00}(Q_2)$ and $f_{yx}(Q_1)$ to have higher magnitude than the other values, Q_2 and Q_1 being the quantization steps of the DC coefficient in the second and first JPEG compressions, respectively, whereas in the absence of NA-JPEG only $f_{00}(Q_2)$ will have higher magnitude. Moreover, in the absence of NA-JPEG the value of $f_{ij}(Q)$ will mainly depend on the image content, so that for a given Q we can assume that $f_{ij}(Q)$ varies very little with (i, j), since the histogram of the DCT coefficients remains quite similar for different shifts.

In order to capture this behavior of $f_{ij}(Q)$ we define the integer periodicity map (IPM) at the quantization step Q as

$$M_{ij}(Q) \triangleq \frac{|f_{ij}(Q)|}{\sum_{i'j'} |f_{i'j'}(Q)|}, \quad 0 \le i \le 7, 0 \le j \le 7. \quad (4)$$

In the case of a NA-JPEG image $\mathbf{M}(Q_1)$ will have a single entry much greater than the others at the location (y, x) corresponding to the shift of the primary compression, whereas in the absence of NA-JPEG $\mathbf{M}(Q)$ will be nearly uniform for every $Q \ne Q_2$. Examples of IPMs for both NA-JPEG and single JPEG compressed images are shown in Fig. 2.

In the proposed algorithm, the uniformity of each IPM is measured by its min-entropy, defined as

$$H_\infty(Q) \triangleq \min_{ij}(-\log M_{ij}(Q)). \qquad (5)$$

A high min-entropy corresponds to a mostly uniform IPM, whereas a IPM with a high peak will be characterized by a
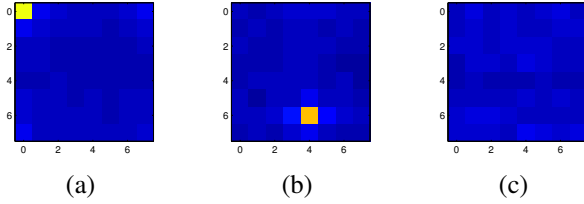
Fig. 2. Examples of IPMs: (a) $Q = Q_2$; (b) $Q = Q_1 \neq Q_2$, $(y,x) = (6,4)$, $H_\infty = 2.56$; (c) $Q \neq Q_2$, absence of NA-JPEG, $H_\infty = 5.23$.

low min-entropy. An image will be classified as NA-JPEG if there is a $Q \neq Q_2$ such that $H_\infty(Q) < T_1$, where T_1 is a suitable threshold, and the relative shift, computed as $(y,x) = \arg\min_{(i,j)} M_{ij}(Q)$, is different from $(0,0)$. When more than one Q satisfies the above condition, the Q achieving the lowest min entropy is selected. The above strategy works well when the quantization step of the primary compression is different from Q_2. When $Q_1 = Q_2$, $\mathbf{M}(Q_2)$ already shows a high peak at $(0,0)$ due to the last compression, which masks the peak due to the first compression.

4. EXPERIMENTAL RESULTS

The image dataset used for experimental validation is composed by 100 non-compressed TIFF images, having heterogeneous contents, coming from three different digital cameras (namely Nikon D90, Canon EOS 450D, Canon EOS 5D) and each acquired at its highest resolution; each test is performed by cropping the central portion using four different image sizes (128×128, 256×256, 512×512, and 1024×1024). For simulating NA-JPEG, each original image is JPEG compressed with a quality factor QF_1, decompressed, cropped by a random shift $(i,j) \neq (0,0)$, with $0 \leq i \leq 7$, $0 \leq j \leq 7$, and JPEG compressed with another quality factor QF_2. The absence of NA-JPEG is simulated by simply compressing the original image with a quality factor QF_2. The quality factor of the first compression (QF_1) is chosen so that the quantization step of the DC coefficients (Q_1) ranges from 2 to 16 with step 1, whereas the quality factor of the second compression (QF_2) is chosen so that the quantization step of the DC coefficients (Q_2) ranges from 1 to 16 with step 1. The case $Q_1 = 1$ is avoided because it is undetectable with the proposed method. This results in 240 possible combinations of (QF_1, QF_2) for each tampered image and 16 different QF_2 for each original image, yielding a total of 24000 tampered images and 1600 original images for each of the four sizes.

For each test, we recorded the minimum value of $H_\infty(Q)$ such that $(y,x) \neq (0,0)$. This allows us to simulate the behavior of the proposed detector for several choices of T_1. In Fig. 3, we show the maximum accuracy of the detector for different values of QF_2 and different sizes. The accuracy is
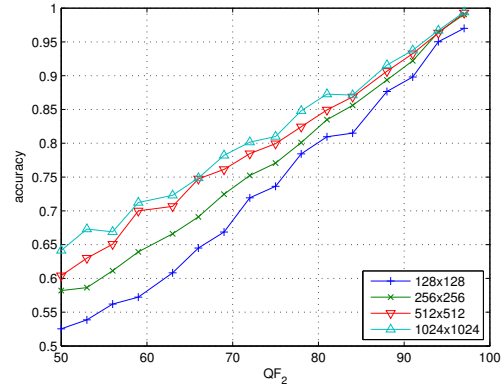
Fig. 3. Accuracy of the proposed detector for different JPEG qualities QF_2 and different image sizes.

Table 1. Maximum accuracy (%) for image size 1024×1024.

QF_1 \ QF_2	50-57	58-67	68-76	77-85	86-95	96
50-57	76.3	95.1	96.6	97.2	98.0	98.0
58-67	82.9	79.1	97.3	98.7	99.3	99.0
68-76	63.2	80.0	80.4	98.8	99.7	100.0
77-85	54.1	55.7	70.9	79.7	99.9	100.0
86-95	54.1	54.1	53.7	57.7	73.2	100.0

evaluated as the percentage of correctly classified images, averaged over all possible QF_1 values. Compared to [4], our detector is about 3% more accurate for similar image sizes. It is worth noting that for $QF_2 > 70$ our detector needs only a 256×256 image to achieve the best performance of [4], which was obtained using 1600×1200 images.

In Table 1, we show the maximum accuracy of the detector for different combinations of QF_1 and QF_2, when the image size is 1024×1024, whereas in Table 2 we show the probability of detecting NA-JPEG for a probability of false alarm equal to zero. The above results show that when $QF_2 - QF_1 > 10$ NA-JPEG is detected with very high accuracy ($> 95\%$), while it is still detected with about 80% accuracy when $QF_1 - QF_2 < 10$ and $QF_1 < 76$. The fact that the performance decreases sharply when QF_1 is similar to QF_2 depends on the fact that the proposed algorithm can not detect NA-JPEG when $Q_1 = Q_2$. Noticeably, the proposed method maintains a very high probability of detection even if we force the probability of false alarm to zero, confirming the reliability of the min-entropy feature.

In order to evaluate the ability of the algorithm to estimate the correct quantization step Q_1 and the correct shift (y,x), we also performed a test by setting the thresholds so as to achieve a probability of false alarm equal to zero and we recorded the estimated parameters. The accuracy, defined as the percentage of correctly identified Q_1 and (y,x) over the

Table 2. Probability of detection ($\%$) for a probability of false alarm equal to zero, for image size 1024×1024.

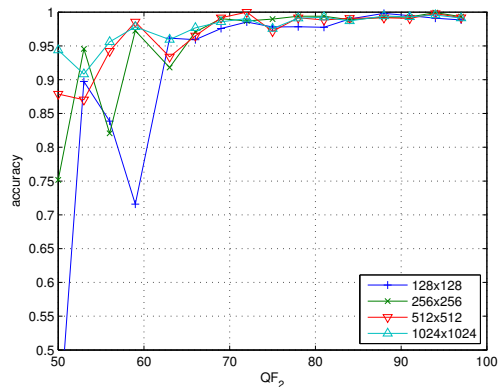| QF_1 \ QF_2	50-57	58-67	68-76	77-85	86-95	96
50-57	50.4	89.8	92.3	93.8	96.0	96.0
58-67	62.7	54.9	94.1	96.8	98.6	98.0
68-76	18.7	55.6	55.9	97.2	99.3	100.0
77-85	1.6	5.4	33.3	55.7	99.9	100.0
86-95	1.0	1.0	0.9	4.9	38.0	100.0

images detected as NA-JPEG, is shown in Fig. 4, for different values of QF_2 and different sizes. For $QF_2 > 66$, the proposed method identified the correct Q_1 and (y, x) in over 98% of the images recognized as NA-JPEG.
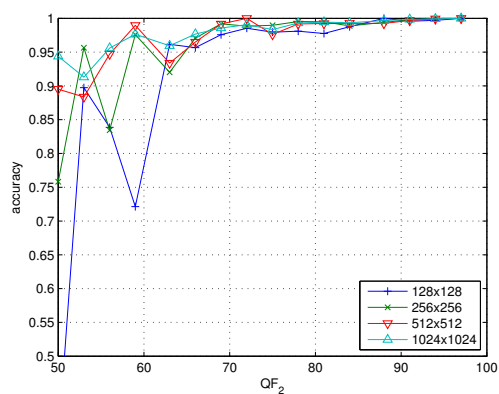
5. CONCLUSIONS

A simple and reliable technique to detect the presence of non-aligned double JPEG compression has been proposed. Differently from previous methods based on SVM classifiers, our approach relies on a detector using a single feature. The method is based on the observation that the DCT coefficients exhibit an integer periodicity when the DCT is computed according to the grid of the primary compression. Such a behavior can be efficiently detected by measuring the non uniformity of a suitably defined integer periodicity map (IPM), in which every entry of the map depends on the DCT statistics for a particular grid shift. The presence of NA-JPEG is detected by applying a threshold detector to the min-entropy of the IPM. Experimental results show that the proposed detector achieves a higher detection accuracy than previously proposed methods and is able to analyze smaller images. Moreover, the proposed method is able to accurately estimate both the quantization step and the grid shift of the primary JPEG compression, which can be used to perform more advanced analyses. As an example, we are currently exploring the possibility of using the estimated parameters to derive a statistical model of DCT coefficients of NA-JPEG images, which can be used for the automatic localization of tampered regions following an approach similar to [3]. Future research will also be devoted to identify NA-JPEG when the second compression uses the same quantization step as the primary compression.

6. REFERENCES

[1] H. Farid, "A survey of image forgery detection," *IEEE Signal Processing Mag.*, vol. 2, no. 26, pp. 16–25, 2009.

[2] H. Farid, "Exposing digital forgeries from JPEG ghosts," *IEEE Trans. on Information Forensics and Security*, vol. 4, no. 1, pp. 154–160, 2009.

(a)

(b)

Fig. 4. Percentage of correctly estimated primary compression parameters: (a) quantization step Q; (b) shift (y, x).

[3] Z. Lina, J. He, X. Tang, and C.-K. Tang, "Fast, automatic and fine-grained tampered JPEG image detection via DCT coefficient analysis," *Pattern Recognition*, vol. 42, no. 11, pp. 2492–2501, Nov. 2009.

[4] W. Luo, Z. Qu, J. Huang, and G. Qui, "A novel method for detecting cropped and recompressed image block," in *Proc. of ICASSP 2007*, 2007, vol. 2, pp. II–217–II–220.

[5] Z. Qu, W. Luo, and J. Huang, "A convolutive mixing model for shifted double JPEG compression with application to passive image authentication," in *Proc. of ICASSP 2008*, 2008, pp. 1661–1664.

[6] M. Barni, A. Costanzo, and L. Sabatini, "Identification of cut & paste tampering by means of double-JPEG detection and image segmentation," in *Proc. of ISCAS 2010*, 2010.

[7] J. Lukáš and J. Fridrich, "Estimation of primary quantization matrix in double compressed JPEG images," in *Digital Forensic Research Workshop*, 2003.