International Journal of Engineering Research and Applications (IJERA) ISSN: 2248-9622
International Conference on Humming Bird (01st March 2014)

RESEARCH ARTICLE OPEN ACCESS

Histogram Analysis of Doubly Compressed JPEG Images for Forgery Detection

Krishna Sobhan , V R Bhuma2

1 Krishna.Sobhan. Author is currently pursuing M.E (Computer science and Engineering) in Vins christian college of Engineering, e-mail:krishna.sobhan@gmail.com.

2 V R Bhuma, the author is working as an Assisstant Professor in Information Technology Department in Vins Christian College of Engineering,Chunkankadai.

Abstract:-
The report is based on a method for estimation of quantization matrix on doubly compressed JPEG images. The characteristic features that occur in DCT histograms of individual coefficients due to double JPEG compression are identified.The double compression detection techniques and quantization matrix estimation are used in the analysis of JPEG files and in digital forensic analysis for detection of digital image forgery.It has been analyzed that by using this method, it is able to fool forensic methods designed to detect evidence of JPEG compression in decoded images, determine an image's origin, detect double JPEG compression, and identifying cut-and-paste image forgeries. It is also shown that how the proper addition of noise to an image's DCT coefficients can sufficiently remove quantization artifacts which act as indicators of JPEG compression while introducing an acceptable level of distortion. Here taking the viewpoint of a forensic analyst show how it is possible to counteract the above said anti-forensic method by revealing the traces of JPEG compression.
Keywords— Digital Forensics,tamper detection,copy- move forgery, double JPEG compression, anti-forensics**.**

I. INTRODUCTION

A double compressed JPEG file is created when a JPEG image is decompressed and then resaved with a different quantization matrix. There are at least two reasons why forensic experts should be interested in double compressed images and the estimation of the primary quantization table[2]. First, double compressed JPEG images often result from digital manipulation (forgeries) when a portion of the manipulated image is replaced with another portion from another image and resaved. In this case, the pasted portion will likely exhibit traces of only a single compression while the rest of the image will exhibit signs of double compression. This analysis [4]could in principle be used to identify manipulated areas in digital images. Second, doubly compressed images are often produced by steganography. For some steganalytic methods, it is very important to estimate the primary quantization matrix to facilitate accurate and reliable steganalysis.

Several image forensic techniques generate the statistical footprints left by JPEG compression. When an image is compressed using JPEG, the histogram of the quantized discrete cosine transform (DCT) coefficients exhibits a characteristic comb-like shape[7]. This fact has been employed to find the original quantization matrix used to compress the image, to identify double JPEG compression and copy-move forgeries. It have shown that adding noise with a certain distribution to the quantized DCT coefficients is sufficient to remove the statistical traces left by JPEG compression and regenerate the original coefficient distribution. However, the dithering signal added to destroy the JPEG compression footprints leaves traces in the tampered image. This anti-forensic tool[3] effectively restores the original distribution of DCT coefficients, but it cannot recover the underlying image content lost during quantization. Therefore, it results in an overall degradation of the original image quality.

The main objective of the paper is to analyze the cost of anti-forensic methods used to remove the traces of JPEG compression. The cost is measured in terms of introduced distortion and loss of image quality. Specifically, it specifies[5] two contributions. First, it analyze the dependency of the mean square error distortion introduced by anti-forensic dithering in terms of the quantization step size and the distribution of the original DCT coefficients. This analysis[8] enables the characterization of the footprint left by the anti-forensic technique in the DCT domain. To support this analysis considering a variation of the algorithm in [7], which makes use of a content dependent perceptual model to add dithering signal mostly in regions of the image. It has been shown that, even in this situation, the forgery is not adequately concealed[12].

Here the focus is on the observation that the anti-forensic dither is a noisy signal which cannot replace content of the image lost during quantization. This introduces visible distortion[10] in the attacked image, which appears as a characteristic grainy noise that allows to discriminate attacked images from original uncompressed images. In the previous work[1] it have analyzed that these traces in terms of distortion are introduced in the tampered image. To this end, the previous work has been extended[1] to the more challenging scenario in which the quantization matrix template is concealed to the forensic analyst. The aim of the proposed detector consist of recompressing the questioned image by varying the coding conditions and observing the amount of grainy noise left by the adversary[6].The results indicate that removing JPEG compression footprints[4] is not a simple technique[1],since the process of footprint removal inevitably introduces new traces in the doctored image.

II.BACKGROUND

A. ANTI-JPEG COMPRESSION FORENSICS

Given a bitmap image that has been JPEG compressed before, anti-JPEG compression forensics[19] aim to reveal the statistical traces left by previous JPEG compression so that the current forensic methods fail to detect resultant images. As specified in the introduction, the two obvious traces introduced by lossy image JPEG compression are the blocking artifacts presented in the spatial domain and the quantization artifacts presented in the DCT frequency domain.The dequantized DCT coefficients after JPEG compression[16] will appear at the multiples of the quantization step.

To remove such quantization artifacts, the method firstly estimates the distribution of an image's transform coefficients before compression distribution[12] , and then inserts anti-forensic dither to the transform coefficients of a compressed image for comparison. As a result of this, the DCT coefficients will spread over the integers than just occur at the multiples of the quantization step[16], which means that the quantization artifacts will be decreased.

The experimental results from [22] show that it can significantly reduce the detection performance of the forensic work.The authors described two anti-JPEG compression techniques. The first method just adds anti-forensic dither to the DCT coefficients for removing quantization artifacts; the second one combines the dither operation and the method of removing blocking artifacts with the purpose of hiding the JPEG compression artifacts..

B. COPY-MOVE FORGERY

Because of the complexity of the problem and its largely unexplored character, the authors[20] thought that the research should start with classifying forgeries by their mechanism, starting with the simple ones, and analyzing each forgery type individually.The first step towards building the tool is taken by identifying one very common class of forgeries, the Copy-Move forgery[17], and developing efficient algorithms for the detection of copy-move forgery.

In a copy-move forgery, a part of the image is copied and pasted into another part of the same image[13]. The technique is done with the intention to make an object is possible whether disappear from the image by covering it with a segment copied from another part of the image. Textured areas, such as grass, foliage, gravel, or fabric with irregular patterns, are ideal for this purpose because the copied areas will likely blend with the background and the human eye cannot easily identify any suspicious artifacts.To make the forgery even difficult to detect[18], then it can use the feathered crop or the retouch tool to further mask any traces of the copied-and-moved segments.

C. ANTI-FORENSIC DITHER

The insertion of the anti-forensic dither corresponds to inject a noise-like signal in the pixel domain[20]. As a result of this, the dithered image is tampered with respect to the doubly JPEG-compressed image. Here for this section, the characterization of analytically the distortion in the DCT domain, showing that it as a function of both the distribution of the original transform coefficients and the quantization step size.Thus it arrived at the conclusion that the energy of anti-forensic dithering is concentrated in the middle DCT frequencies, thus resulting in a grainy noise in the spatial domain. Next, analyze the effect of requantizing the dithered coefficients[23] in a DCT subband using different quantization step sizes.

From the analysis it has been identified that requantizing the dithered coefficients with the original JPEG quantization step annihilates completely the anti-forensic noise.In the JPEG compression standard, a greyscale image is first divided into non-overlapping pixel blocks of size 8 × 8[1]. Then, the DCT of each block is computed. However, the DCT coefficient values typically remain tightly clustered around integer multiples of qi, thus revealing that a] a quantization process has occurred and b] which was the original quantization step.

III. ALGORITHM DESCRIPTION

The questioned image is typically the only source of information available to the forensic analyst. Therefore, forensic techniques[18] analyze the image content in order to find the traces left by specific acquisition, coding or editing operations, which could be of malicious tampering. This fact enables several forensic analysis tasks, including the identification of which camera took a picture, or the detection of double JPEG compression. The input image is broken into 8*8 blocks of pixels. Working from left to right, top to bottom, DCT is applied to each block. DCT is calculated for each entry. An image block with a lot of change in frequency has a very random looking resulting matrix[5]. The block matrix consisting of 64 coefficients where top-left lower frequencies and right bottom have higher frequencies.DCT have is calculated using the equation.

$$D(i,j)=1/4C(i)C(j)\sum_{x=0}^{7}\sum_{y=0}^{7}p(x,y)\cos[(2x+1)i/16]\cos[(2y+1)j/16] \quad (1)$$

The 8*8 block results from equation (1) is usually in the below matrix for JPEG images[2]. The 8*8 block of DCT coefficients by equation (1) is now ready for quantization. The specific quantization matrix is identified when user selects the target quality factor Q in case of JPEG. The forensic analyst generate 8*8 quantization matrix **Qa.Q**a=**Q**a(Qa),where the subscript 'a' refers to the quality factor used by the analyst. The analyst recompresses the image using different quality factor.

The total variation TV (Qa) =TV(**Q**a(Qa)) is computed. With quality factor 50 the matrix renders the below matrix.

Cij=round(Dij/Qij) (2)

To achieve lossy compression, a JPEG encoder quantizes each discrete cosine transform (DCT) coefficient of an image to multiples of a quantization. Depending on the specific frequency and channel, each DCT coefficient, is then quantized by an amount. The full quantization is specified as a table of values a set of values associated with each frequency,n step size, specified by the JPEG quantization matrix[4]. The quantization matrices are compared and is detected.

Example:

$$C=\begin{bmatrix}10 & -9 & 5 & -7 & 4 & 3 & -2 & 4\\ 3 & 7 & 7 & 6 & 5 & 6 & -1 & 5\\ 5 & 4 & 3 & 5 & 3 & -5 & 0 & 7\\ 6 & 2 & 2 & 24 & 3 & 3 & 0 & 2\\ 2 & 0 & 0 & 0 & 2 & 2 & 0 & 8\\ 6 & 7 & 9 & 23 & 9 & 0 & 6 & 3\\ 0 & 1 & 8 & -1 & 7 & 3 & 4 & -9\\ 9 & 8 & 4 & 6 & 6 & 1 & 3 & 0\end{bmatrix}$$

The resultant matrix from equation (2) is again multiplied with the matrix values to get the DCT matrix. This is called inverse DCT. The Inverse DCT have slight difference, since a rounding operation was performed. The main aim is to analyze the cost of anti-forensic methods[5] used to remove statistical traces of JPEG compression.Specifically, it describes two contributions. First, it analyzes the dependency of the mean square error distortion[6] introduced by anti-forensic dithering and the distribution of the original DCT coefficients. Here it is analysed for the evaluation of the cost in term of perceptual quality loss for both the baseline anti-forensic method[8] and the perceptually modified version are described. The comparison here is with respect to the original, uncompressed image.

IV. EXPERIMENTAL RESULTS

For the case of known quantization matrix template[6], let the threshold vary to trace the receiver characteristic curve.Here, the true positive rate is the fraction of JPEG compressed images that were correctly reported to be compressed and the false positive rate is the fraction of uncompressed images that were reported to be compressed.Thus, it does not reveal the performance of the proposed method for different values of the quality factor . In order to observe this each curve that is obtained by considering a subset of the original dataset, which is constructed by taking all the images that were JPEG compressed at quality factor[8], and an equivalent number of uncompressed images selected at random, so as to obtain a balanced image dataset.

All test images were originally obtained using different digital cameras as JPEGs, only two of them were original.To remove JPEG artifacts[9], have resized JPEG images to 83% of their original size using PaintShop Pro 7 and saved them as BMPs. All images were also converted to grayscale. It have continued in the experiments on double compressed images. To prepare the test images, there have been used standard quantization matrices corresponding to quality factors 61, 65, 70, 75, 79, 84, 88, 90, and 95 for both primary and secondary quantization matrices. It have also included cases when the primary quantization matrix was non-standard. Using

all possible combinations of these quantization matrices[2], have prepared the total of 900 different JPEG files.

TABLE I: DETECTION ACCURACY AVERAGED OVER ALL JPEG CODING CONDITIONS

Detector of known matrix template(matrix template : concatenation majority voting)	Accuracy	Detector of unknown matrix template(existing system)	Accurac-y
(6,3)	0.76	(6,3)	0.65
(7,3)	0.87	(7,3)	0.78
(3,9)	0.85	(3,9)	0.81
(4,2)	0.87	(4,2)	0.82
(6,5)	0.83	(6,5)	0.80
(8,2)	0.81	(8,2)	0.76
(2,1)	0.86	(2,1)	0.75
(5,3)	0.83	(5,3)	0.75
(8,4)	0.81	(8,4)	0.72
(7,1)	0.89	(7,1)	0.85
(5,4)	0.83	(5,4)	0.80

The method is classified for DCT coefficients in different files. Thus, the error rate was low because out of the examined DCT coefficients, only few were misclassified. Table I shows the number of misclassified cases for different combinations of quality factors.Table I also shows that errors are more likely to occur when the primary quantization step is followed by a large secondary step, i.e., when the double compression decreases the image quality.

V. CONCLUSION

It is possible to observe that method can be effectively adopted to reveal the traces of JPEG compression anti-forensics, achieving very good results, comparable to the method, for a wide range of quality factors. However, since the proposed method is specifically tailored to detect JPEG compression[1] in the presence of anti-forensics.The proposed method is able to estimate the underlying JPEG quality factor or some elements of the quantization matrix[19], when JPEG compression is detected.Future research will investigate the problem of compression anti-forensics in the field of video coding. The motion-compensation provides a further element both the forensic analyst and the adversary can play with.

REFERENCES

[1] GiuseeppeValenzise ,StefanoTurbaro and Macro Tagliasacchi ,"Revealing the Traces of JPEGCompression Anti-Forensics"in IEEE Transactions on Information Forensics and Security ,February 2013.

[2] G. Valenzise, M. Tagliasacchi, and S. Tubaro, "The cost of JPEG compression anti-forensics," in Proc. Int. Conf. Acoustics, Speech, and Signal Processing, Prague, Czech Republic, May 2011.

[3] A. C. Popescu and H. Farid, "Exposing digital forgeries by detecting traces of resampling," IEEE Trans. Signal Process., vol. 53, no. 2, pp. 758–767, Feb. 2005.

[4] J. Fridrich, D. Soukal, and J. Lukás, "Detection of copy-move forgery," in Proc. Digital Forensic Research Workshop, Cleveland, OH, Aug. 2003.

[5] S. Bayram, H. T. Sencar, and N. Memon, "A survey of copy-move forgery detection techniques," in Proc. IEEE Western New York Image Processing Workshop, Rochester, NY, Oct. 2008.

[6] A. Bianchi, T. De Rosa, and A. Piva, "Improved DCT coefficient analysis for forgery," in Proc. Int. Conf. Acoustics, Speech, and Signal Processing, Prague, Czech Republic, May 2011.

[7] M.C.Stamm , S.K.Tjoa , W.S.Lin and K.J.R.Liu ," Anti-forensics of JPEG compression," in Proc. Int. Conf. Acoustics, Speech, and Signal Processing, Dallas, TX, Apr. 2010.

[8] M. C. Stamm , S. K. Tjoa , W. S. Lin, and K. J. R. Liu, "Undetectable image tampering through JPEG compression anti-forensics," in Proc. Int. Conf. Image Process., Hong Kong, Sep. 2010, pp. 2109–2112.

[9] H. Farid, "Exposing digital forgeries from JPEG ghosts," IEEE Trans. Inf. Forensics Security, vol. 4, no. 1, pp. 154–160, Mar. 2009.

[10] S. Y. Lai and R. Böhme, "Countering counter-forensics: The case of JPEG compression," in Information Hiding. New York: Springer, 2011, pp. 285–298.

[11] G. Schaefer and M. Stich, "UCID" in Proc. SPIE: Storage and Retrieval Methods and

Appli- cations for Multimedia, 2004, vol. 5307, pp. 472–480.

[12] T. Bianchi and A. Piva, "Image forgery localization," IEEE Trans. Inf. Forensics Security, vol. 7, no. 3, pp. 1003–1017, Jun. 2012.

[13] S. Y. Lai and R. Böhme, "Countering counter-forensics: The case of JPEG compression," in Information Hiding. New York: Springer,2011, pp. 285–298.

[14] H. Farid, "Exposing digital forgeries from JPEG ghosts," IEEE Trans.Inf. Forensics Security, vol. 4, no. 1, pp. 154–160, Mar. 2009.

[15] P. Bestagini, A. Allam, S. Milani, M. Tagliasacchi, and S. Tubaro,"Video codec identification," inProc. Mar. 25–30, 2012, pp. 2257–2260.

[16] J. Fridrich, D. Soukal, and J. Lukás, "Detection of copy-move forgery," in Proc. Digital Forensic Research Workshop,Cleveland, OH, Aug. 2003.

[17] S. Bayram, H. T. Sencar, and N. Memon, " Copy-moveforgery detection techniques," in Proc. IEEE Western New York Image Processing Workshop, Rochester, NY, Oct. 2008.

[18] A. Bianchi, T. De Rosa, and A. Piva, "Improved DCT coefficient analysis," in Proc. Int. Conf.Acoustics, Speech, and Signal Processing, Prague, Czech Republic,May 2011.

[19] M. C. Stamm, S. K. Tjoa, W. S. Lin, and K. J. R. Liu, "Undetectable image tampering," in Proc.Int. Conf. Image Process., Hong Kong, Sep. 2010, pp. 2109–2112.

[20] W. Luo, Y. Wang, and J. Huang, "Security analysis on steganography for JPEG decompressed images," IEEE Signal Process. Lett., vol. 18, no. 1, pp. 39–42, Jan. 2011.

[21] H. Farid, "Exposing digital forgeries from JPEG ghosts," IEEE Trans.Inf. Forensics Security, vol. 4, no. 1, pp. 154–160, Mar. 2009.

[22] P. Bestagini, A. Allam, S. Milani, M. Tagliasacchi, and S. Tubaro,"Video codec identification," inProc. Int. Conf. Acoustics, Speech, andSignal Processing, Mar. 25–30, 2012, pp. 2257–2260.

[23] S. Y. Lai and R. Böhme, "Countering counter-forensics," in Information Hiding. New York: Springer,2011, pp. 285–298.

[24] T. Pevny, P. Bas, and J. Fridrich, "Steganalysis by subtractive pixel adjacency matrix," IEEE Trans. Inf. Forensics Security, vol. 5, no. 2,pp. 215–224, Jun. 2010.