ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 4, Issue 8, August 2015

Detection of Double JPEG Compression on Color Images using EBSF Method

Ann Josy John[1], Jissy Liss Jose[2]

Student, Computer Science, Mangalam College of Engineering & Technology, Kottayam, India[1]

Asst.Professor, Computer Science, Mangalam College of Engineering & Technology, Kottayam, India[2]

Abstract: Detection of double jpeg compression with in an image with the same quantization matrix is a challenging problem. Detection of double JPEG compression plays an important role in digital image forensics. Some successful approaches have been proposed to detect double JPEG compression when the primary and secondary compressions have different quantization matrices. However, detecting double JPEG compression with the same quantization matrix is still a challenging problem. Here, an effective error based statistical feature extraction scheme is presented to solve this problem. First, a given JPEG file is decompressed to form a reconstructed image. An error image is obtained by computing the differences between the inverse discrete cosine transform coefficients and pixel values in the reconstructed image. Two classes of blocks in the error image, namely, rounding error block and truncation error block, are analyzed. Then, a set of features is proposed to characterize the statistical differences of the error blocks between single and double JPEG compressions. Finally, the K nearest neighbour classifier is employed to identify whether a given JPEG image is doubly compressed or not.

Keywords: Double JPEG compression, rounding error, truncation error,,color image.

I. INTRODUCTION

Verifying the integrity of digital images is an important research area of image processing.In this digital world there is a need of digital images in all the areas like medical imaging, surveillance systems,forencics[2-7] etc.in order to store these images its size should be reduced ,for that compression techniques like jpeg are used. In this work, here focuses on detecting double compression on color images. When altering an JPEG image, it is loaded into a photo editing software (decompressed) and changes its certain portions, the image is then resaved then it become compressed again. The quantization matrix of the orginal image is called as primary quantization matrix. The quantization matrix of the resaved image is called as secondary quantization matrix. If the primary and secondary quantization matrix are identical, then the resaving (double compressing) operation brings into the image specific changes.

Detecting these changes plays a valuable role in identifying image forgeries. Detecting the traces of double compression also is helpful in other research fields such as steganography . The jpeg compressions can be again classified into aligned jpeg compression and non-aligned jpeg compression.The aligned jpeg compressions are again classified into aligned compression with same quantization matrix and with different quantization matrix.some succesfull methods are proposed earlier to detect
the compressions Farid [9] proposed to measure the underlying periodic artifacts of JPEG coefficient histograms via Fourier transform for detection of double JPEG compression. In [10], Fu et al. found that the distribution of the first digits of JPEG alternating-current (AC) coefficients follows a generalized Benford's law, which can be used to distinguish between singly and doubly JPEG compressed images. Moreover, in [11], by applying the generalized Benford's law to some specially selected individual AC modes, the performance of detecting double JPEG compression can be further improved. Benford's law is also used to estimate the number of JPEG compressions [22],all these methods require to convert the images into gray scale value.

The above mentioned methods are not suitable for detecting the double jpeg compression on color images. Extraction of image features on all the color planes is very difficult.The col or is a perception produced by the stimulation of rectinal cells by light. The image usually lies in a plane YCbCr, RGB color spaces. Most of the algorithm result in works only on gray scale images and it might not correctly extract the image features.The proposed EBSF method first apply Idct on compressed image then identify the error blocks from the given image then compute the features in all the three planes cprovide these as an input to the KNN classifier it will generate the output as single or double compressed image as in Fig 1.

II. BASCIS OF JPEG COMPRESSION

JPEG is an international lossy compression standard from joint photographic expert group. JPEG is a lossy image compression method. It employs a transform coding method using the DCT (Discrete Cosine Transform).

First the image is converted into 8*8 blocks then apply DCT on the image blocks to obtain $d_{i,j}$ these values can then quantized using a quantization matrix.

If these image blocks are manipulated for the case of image splicing it should be recompressed again to make the image doubly compressed. During decompression

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 4, Issue 8, August 2015

these operations are to be performed in reverse to obtain the orginal image The detection of this double compression is focused in this paper.

A. *Errors in jpeg compression*

The jpeg compression mainly consist of mainly are, discrete cosine transform of 8*8 image blocks, quantiztion and entropy encoding. During decompression these operations performed in reverse steps, order: entropy decoding, de-quantization and inverse DCT. During these operations there are three errors may be occurred. First is quantization error which may be occurred during the jpeg compression process it is the difference between the values of divided float valued jpeg coefficients before and after rounding to a certain integer value.The next two errors can be found only during the decompression they are rounding and truncation errors. After perfoming IDCT on dequantized jpeg coefficients float values should be truncated and rounded to nearest integer value. The difference in value between the float IDCT coefficients and the rounded integer value is called rounding error which falls in the range of [0-255].If the rounded value exceeds this range it wiil truncated to fall within this range ie The difference in value between float IDCT coefficients and this limit is called truncation error.

For the detection of image compression details here only rounding and truncation error blocks are considered. During calculation it is found that rounding error blocks have all the values falls in the range of [-0.5-0.5].while truncation error blocks having values such that atleast one value of the coefficients in the block exceeds the limit of [-0.5-0.5].

III. PROPOSED SYSTEM

The detection of double jpeg compression is very important in this technologically advanced world. From the previous works in the same area it is understood that a new scenario is needed to detect the double jpeg compression on colour images. The traditional method's success usually relay only on the grey scale images. The Error based statistical feature extraction method works well on colour images as it is already works for grey scale images.

The EBSF method tries to find out the rounding and truncation error blocks with in the given image and then extract the features from the given error blocks as in Fig 1The number of error blocks reduces as the number of times the compression is performed ,that is the number of error blocks in single compression is more than that of a double compressed image.

A. Relationship between error blocks during consecutive compressions

The proposed system assumed that the jpeg images are compressed only with the same quantization matrix.According to the process the error image obtained by performing the rounding and truncation operation.The error blocks thus obtained can be denoted as R_{n}.

$$R_n = TR(IDCT(D_n)) - TR(D_n)$$

Where TR is the rounding and truncation operations performed on the given image.IDCT denotes the inverse discrete cosine transform on all the blocks ,D_n denotes the dequantized values with in the image.

$$D_n = K_n \times Q$$

Where Q is the quantization matrix and K_n denotes the coefficient matrix of n times compressed image. × denotes the component wise multiplication on each pixels.

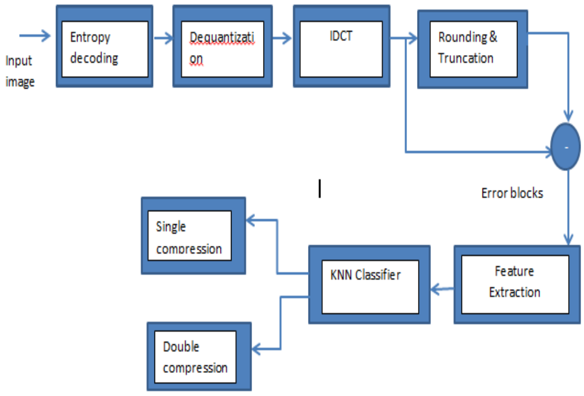

Fig 1: Proposed block diagram

After each compressions there are certain block values doesnot varies and such blocks can be called as stable error blocks can be denoted by M_n ,other blocks whose value $M_n \neq 0$ are called unstable error blocks. While analysing the error blocks for the detection of double compression these stable blocks will not provide any usefull information so it should be eliminated before computations. In order to distinguish n times compressed image from n+1 compressed image there is a need to focused on the error blocks ,at that time stable error blocks will not provide any information and it should be excluded.

Another important idea is that the unstable error blocks wiil not increase with the increase in the number of compressions, ie the unstable error blocks reduced as the number of compression increases.The stable error blocks are always stable which will not changes its value during recompressions .From the above analysis two main ideas should be considered here stable error blocks should be eliminated before computation and the decrease in number of unstable blocks indicates that there are only few compressions are occured .

B. Feature extraction in color images

This subsection deals with the extraction of a set of features in color jpeg images for KNN classifier. The images are choosen from the UCID database which consist of uncompressed and singly compressed images.The statistical difference between single and double compression can be identified from the characteristics of both rounding and truncation error blocks.In the EBSF method there are three sets of features can be used.The first subset is EBSF_spatial extracted directly from the rounding and truncation error blocks ,which are the mean

DOI 10.17148/IJARCCE.2015.4882

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 4, Issue 8, August 2015

and variance of the maximum values of the rounding and truncation error blocks defined as,

$$\text{Mean}(|Rn(i,j)|) = \frac{\sum_{l=1}^{L}\sum_{I=0}^{7}\sum_{J=0}^{7}|Rn(i,j)|}{64\,L}$$

$$\text{Var}(|Rn(i,j)|) = \frac{\sum_{L=1}^{L}\sum_{i=0}^{7}\sum_{j=0}^{7}|Rn(i,j)|-mean|Rn(i,j)|2}{64L}$$

L denotes the number of unstable rounding blocks in the image.The mean Mean (|Tn(i,j)|) and varience var(|Tn(i,j)|) of the truncation error blocks can be also calculated using the same formulae.

The second feature sets based on the DCT coefficients which helps to distinguish between single and double compression.EBSF_dct which needs to calculate Wn, which is obtained by compressing the reconstructed image again and then calculating the de-quantized JPEG coefficient changes between the two consecutive compressions .Mean and Varience of the Wn over AC component of rounding block w_r can be calculated for computing the feature set defined by,

$$\text{Mean}(|w_r(u,v)|) = \frac{\sum_{l=1}^{L}\sum_{(U,V)\neq 0}|(Wr(u,v))|}{63L}$$

$$\text{Var}(|Wr(u,v|)= \frac{\sum_{l=1}^{L}\sum_{U,V\neq 0}(|Wn(i,j)|)-mean(|Wr(u,v))|))2}{63L}$$

Similarly the truncation error block features are also added.

The third set of features EBSF_ratio which will contain only one feature. It is the ratio of the number of unstable rounding error blocks, nr , to the number of all unstable error blocks, na, i.e., calculated by nr /na. So there are thirteen features all together four from EBSF_spatial, eight feature from EBSF_dct and one from EBSF_ratio are provided as an input to train the KNN to find out whether the given image is single or double compressed

IV. Experimental Result

The EBSF method perform the faster and accurate classification of single and double jpeg compressed images. The images in the UCID and NRCS data bases are effectively extracted and identifies the features of rounding and truncation error blocks in all the three colour planes. The KNN classifier is compared to be out perform SVM classifier. The input training set and comparison results in the database is shown in the fi 2.

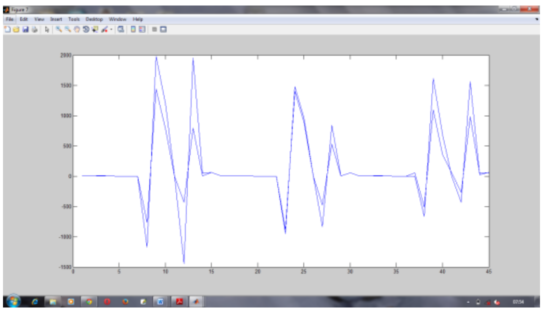

Fig 2: Knn classifier result

and the results after classification of the images features are given in fig 3

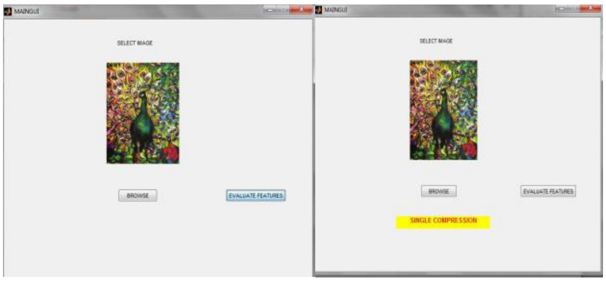

Fig 3: Detection result

The EBSF method can be used also to detect the double compression from the triple compression also.

V. Conclusion

Here a learning based method is there to detect double jpeg compression with the same quantization matrix, which is easy to implement while shows promising performance. First analyzed the error blocks in jpeg compression only the rounding and truncation error can be utilized to discriminate between singly and doubly compressed images with the same quantization matrix. Based on the analysis, error-based statistical features were extracted from rounding and truncation error blocks separately. Finally, with the extracted features, the knn classifier is applied for detecting single or double jpeg compression within the image .

ACKNOWLEDGMENT

The first author would like to thank all those people who guided and supported. Without their valuable guidance and support, this task was not possible and also likes to thank colleagues for their discussions and suggestions.

References

[1] Jianquan Yang, Jin Xie, Guopu Zhu, Senior Member, IEEE,, "An Effective Method for Detecting Double JPEG Compression With the Same Quantization Matrix" ieee transactions on information forensics and security, vol. 9, no. 11, november 2014

[2] r. n. neelamani, r. de queiroz, z. Fan, S. Dash, and R. G. Baraniuk,"JPEG compression history estimation for color images," IEEE Trans.Image Process., vol. 15, no. 6, pp. 1365–1378, Jun. 2006.

[3] W. Luo, J. Huang, and G. Qiu, "JPEG error analysis and its applications to digital image forensics," IEEE Trans. Inf. Forensics Security, vol. 5,no. 3, pp. 480–491, Sep. 2010.

[4] T. Bianchi and A. Piva, "Image forgery localization via block-grained analysis of JPEG artifacts," IEEE Trans. Inf. Forensics Security, vol. 7, no. 3, pp. 1003–1017, Jun. 2012.

[5] H. Farid, "Exposing digital forgeries from JPEG ghosts," IEEE Trans.Inf. Forensics Security, vol. 4, no. 1, pp. 154–160, Mar. 2009. Z. Lin, J. He, X. Tang, and C.-K. Tang, "Fast, automatic and finegrained

[6] aAmpered JPEG image detection via DCT coefficient analysis," Pattern Recognit., vol. 42, no. 11, pp. 2492–2501, Nov. 2009.

[7] J. Luká˘s and J. Fridrich, "Estimation of primary quantization matrixin double compressed JPEG images," in Proc. Digit. Forensic Res. Workshop, 2003, pp. 5–8.

[8] A. C. Popescu and H. Farid, "Statistical tools for digital forensics," inProc. 6th Int. Workshop Inf. Hiding, 2004, pp. 128–147.

[9] D. Fu, Y. Q. Shi, and W. Su, "A generalized Benford's law for JPEG coefficients and its applications in image forensics," Proc. SPIE, truncation error blocks vol. 6505, p. 65051L, Feb. 2007. B. Li, Y. Q. Shi, and J. Huang, "Detecting doubly compressed features"

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 4, Issue 8, August 2015

[10] T. Pevný and J. Fridrich, "Detection of double-compression in JPEG images for applications in steganography," IEEE Trans. Inf. ForensicsSecurity, vol. 3, no. 2, pp. 247–258, Jun. 2008.
[11] C. Chen, Y. Q. Shi, and W. Su, "A machine learning based scheme for double JPEG compression detection," in Proc. 19th Int. Conf. Pattern Recognit., Dec. 2008, pp. 1–4.
[12] A. Westfeld, "F5—A steganographic algorithm: High capacity despite better steganalysis," in Proc. 4th Int. Workshop Inf. Hiding, 2001, pp. 289–302.
[13] N. Provos, "Defending against statistical steganalysis," in Proc. 10th USENIX Secur. Symp., Aug. 2001, pp. 323–336.
[14] [16] F. Huang, J. Huang, and Y. Q. Shi, "Detecting double JPEG compressionwith the same quantization matrix," IEEE Trans. Inf. Forensics Security,vol. 5, no. 4, pp. 848–856, Dec. 2010.
[15] S. Lai and R. Böhme, "Block convergence in repeated transform coding: JPEG-100 forensics, carbon dating, and tamper detection,"in
[16] Proc. IEEE Int. Conf. Acoust., Speech, Signal Process., May 2013, pp. 3028–3032.

BIOGRAPHIES

Ann Josy John Department of Computer Science & Engineering,Mangalam college of Engineering, Ettumanoor, Kerala, India.

Jissy Liss Jose, Assistant Professor, Department of Computer Science and Engineering, Mangalam College of Engineering, Ettumanoor, Kerala, India.