Detection of JPEG Ghost in Non-Aligned Spliced Region of JPEG Images

R K Khatri
Scientist,
Defence Research & Development Organisation
Jodhpur, India

Rajesh Purohit
Associate Professor,
MBM Engineering College
Jodhpur, India

***Abstract*— JPEG Ghost is format based image forensics analysis technique to expose forgery due to image splicing. When the copied region initially has lower compression quality than source image, this technique is very effective, easy to use and results can be visually verified. Many algorithms have been proposed based on this approach for aligned grids. This paper shows, how JPEG ghost is affected and can be recovered when there is misalignment of grid in spliced region.**

Keywords*— DCT, Digital Forensics, Double Compression, Image Splicing, JPEG Ghost

NOMENCLATURE

BACM – Blocking Artifacts Characteristics Matrix, DCT – Discrete Cosine Transform, IDCT – Inverse DCT, SVM – Support Vector Machine, UCID – Uncompressed Color Image Database.

I. INTRODUCTION

In the current boom of digitalization, it is getting difficult to visually detect the tampering of images due to high availability of sophisticated image editing software. Therefore requirement of software tools and techniques is ever increasing for authentication of visual information. To increase trust in images, the field of digital image forensics has emerged as candidate of active research in digital content creation community. One solution is to use digital watermarking at the time of image creation. It is active technique for image authentication, but requires specially equipped camera system at the time of recording. However, most of images available are taken from normal digital cameras, therefore this technique cannot be used on existing images. In contrast to this, more practical solution is to use passive techniques, that operates in the absence of pre-embedded watermark or signature. It works on the assumption that digital forgeries may not be perceived visually but change the underlying statistical pattern of an image. This alteration or artifacts can be used as clue for exposing the tampering in images. In [8], digital image forensic techniques have been broadly classified into five categories: 1) pixel-based techniques; 2) format-based techniques; 3) camera-based techniques; 4) physics based techniques; and 5) geometric-based techniques. JPEG is most prevailing and preferred format for storage and sharing. Due to high availability of image content in this format, lot of literature is available on JPEG based forensics analysis. More about JPEG format based image forensics techniques is discussed in the next section.

II. JPEG FORENSICS

In attempt to forge an image in JPEG format, artifacts arises due to lossy nature of its compression scheme. Therefore, format based technique is widely used to expose traces of tampering in JPEG images. In this technique the pattern or artifacts generated due to multiple encoding and decoding steps are analyzed to expose the traces of tampering.

A. JPEG Compression Scheme

JPEG standard specifies two compression schemes: a lossless predictive scheme and a lossy scheme based on DCT. The most popular and lossy compression scheme is baseline method due to its high compression capability with little compromise on quality [14]. The baseline method involves three basic steps as described in [1] and summarized in next para.

1. DCT: In this step, raw image is converted from RGB to YCbCr format, divided into 8 x 8 pixel blocks in raster scan order, shifted from pixel intensity range [0,255] to [–128,127] and DCT coefficients are computed for each 8 x 8 block separately.

2. Quantization: The DCT coefficients obtained in previous steps are divided by quantization tables, and rounded off to the nearest integer.

3. Entropy Encoding: This step involves lossless bit level compression that transforms the quantized DCT coefficients into a bit stream of compressed data using Huffman coding. The decoding of a compressed data stream involves the inverse of the previous three steps, taken in reverse order i.e. entropy decoding, de-quantization, inverse DCT and conversion to RGB color space for display of image.

B. Error Propagation in Compression

DCT-based JPEG compression method is lossy in nature and introduce some errors in different stages of encoding and decoding. There are three different kinds of errors described in [5], that are introduced into the compression and decompression process. The first one is the quantization error, which creeps in the compression process. In this process image is divided into 8 x 8 blocks of pixels and DCT is applied to get the coefficients. The obtained DCT coefficients are then divided by integer values defined in quantization tables and then rounded to nearest integer. The difference between the actual float values of the divided DCT coefficients and the rounded integer values is called quantization error. The second and third kinds of errors both exist in the decompression process. After applying IDCT to the de-quantized JPEG coefficients, the obtained float values are truncated in range [0 255]. The values above 255 or below 0 are truncated to 255 and 0 respectively and this step introduce truncation error. In addition, the float values in range 0 to 255 are rounded to the nearest integers while reconstructing the image in the spatial domain. This step introduces the rounding error.

III. JPEG ARTIFACTS

The various JPEG artifacts introduced due to splicing has been categorized in [11], as Double Compression, Blocking Artifacts and JPEG Ghost.

A. Double Compression

Any digital manipulation requires an image to be loaded into a photo-editing software and resaved after modifications. Since most images are stored in JPEG format, it is likely that both the original and manipulated images are stored in this format. In this scenario manipulated image is compressed twice. Because of the three kinds of errors mentioned in previous section, i.e. quantization error, truncation error, and rounding error, the doubly compressed image may have some difference from the original one even if the image is doubly compressed with the same quantization. This double compression introduces specific artifacts not present in singly compressed images. The presence of these artifacts can therefore be used as evidence of some manipulation [1].

B. Blocking Artifacts

In the JPEG encoder, the image is first divided into 8 x 8 non-overlapping blocks. Each block is DCT-transformed, quantized and then entropy encoded to yields a data stream. When there is no compression, pixel differences across blocks should be similar to those within blocks. If the image is JPEG-compressed, the differences across blocks should be different due to blocking artifacts [10]. If image splicing is done, the pattern of blocking artifacts is disturbed and this can be used as clue to detect the tampering using BACM as described in [13].

C. JPEG Ghost

Every time a JPEG image is re-compressed, the statistics of its compression coefficients slightly change. JPEG ghost detection technique first proposed in [7], makes use of artifacts introduced in image due to 8 x 8 blocks, DCT transform and double quantization. It displays localized re-compressed forged region at different contrast than un-forged image, which appears in resultant difference images after subtracting it from its various re-compressed version at different quality levels.

IV. RELATED WORK

When creating a digital forgery, it is often necessary to combine several images. JPEG ghost principle describe an image splice detection technique where the part of an image is initially compressed at a lower quality than the rest of the image. Ghost effect can be shown visually as well as quantitatively. Six different papers have been explored, in which authors proposed image splicing detection based on JPEG ghost principle. In [3], a feature vector is calculated using error difference in images, to train SVM classifier. Doctored parts in image can also be located by examining the double quantization effect hidden among the DCT coefficients as described in [9]. Detection of double quantization with improved accuracy, is proposed in [12] based on training classifier, developed using CASIA tampered and authentic image dataset. The result of JPEG ghost can be visually verified in small number of cases. An attempt has been made for automation of JPEG ghost detection using graph based segmentation in [2] and training classifier based approach in [4]. All these methods proposed in literature are based on direct measurement of error variation at different stages, due to JPEG compression and de-compression process.

V. PROPOSED METHOD

All methods discussed above have shown test results assuming aligned grid splicing. Work proposed in this paper expose JPEG ghost for both aligned and non-aligned spliced (copied) region of JPEG images. When image splicing is done, it is very likely (63/64 probability), that the 8 x 8 grid is misaligned. Due to misalignment of grids, JPEG ghost effects is destroyed, and spliced region cannot be visually or quantitatively detected by method proposed in [7]. To recover the ghost effect, all 64 combination of difference images has been analyzed by shifting 8 x 8 block grid alignment in horizontal and vertical direction.

To test the detection of JPEG ghosts, uncompressed TIFF images have been taken from the UCID database [6]. These color images are each of size 512 x 384 and cover a wide range of indoor and outdoor scenes. A selected TIFF image is first saved in JPEG format at qualities q1 and q2, where q1 > q2. Then portion of image at quality q2 is inserted into image of quality q1 and whole image is resaved at quality q1. The Matlab function 'imwrite' has been used to save images in the JPEG format at desired qualities by selection of compression quality in range of [1 100].

VI. RESULTS

One image in TIFF format is selected from UCID database, as shown in Fig. 1.

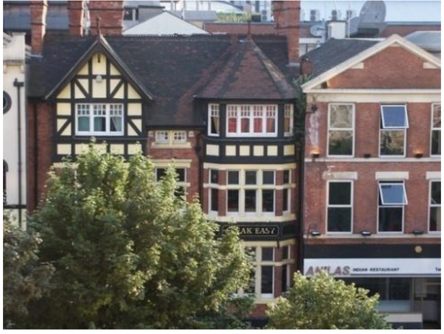

Fig. 1. Image from UCID Database

First image is saved at two JPEG qualities $q_1 = 90$ and $q_2 = 70$. Then a rectangular block of size 200 x 150 pixels is spliced from lower quality JPEG image to higher quality JPEG image and spliced image is saved at quality 90. This is test case for aligned JPEG splicing. Now this images is resaved at different JPEG quality and RGB color difference is computed for each pixel. The resulted difference images are shown in Fig. 2.

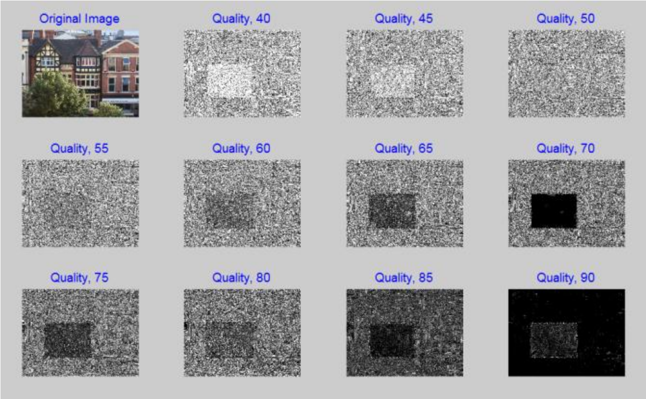

Fig. 2. Display of JPEG Ghost for Aligned Grids

Difference images, shows minimum intensity values i.e. dark region where splicing was done. This is known as JPEG ghost effect. When block is misaligned by shifting spliced region to one pixel right and on pixel bottom, the JPEG ghost effect is destroyed as shown in Fig. 3.

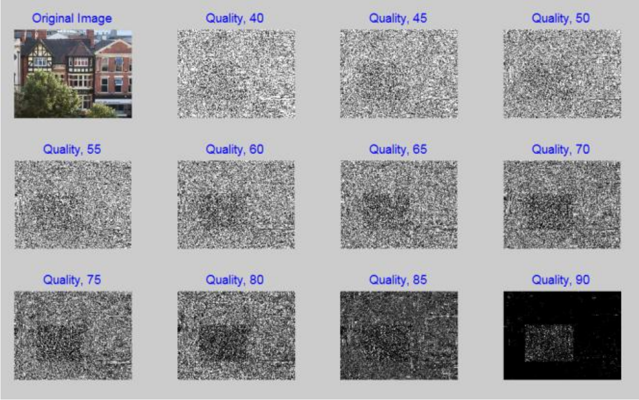

Fig. 3. Display of JPEG Ghost for Mis-aligned Grids

To recover the JPEG ghost, all 64 combination of rows and columns misalignment has been tested, which are possible in 8 x 8 grid. The result of JPEG ghost effect by shifting the image by one pixel right and one pixel bottom is shown in Fig. 4. Here again dark region with minimum intensity values can be seen at quality 70.

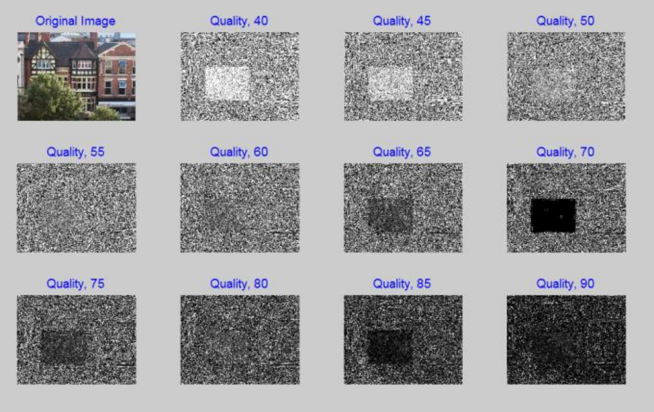

Fig. 4. Display of JPEG Ghost for Re-aligned Grids

VII. CONCLUSION

In this paper, proposed tamper detection algorithm has been tested by splicing rectangular region, where both source and target images having different JPEG qualities. The algorithm can be visually verified by display of Ghost images. It works on misaligned grid and does not require training classifier. However the method fails when source and target images are saved at same compression quality.

REFERENCES

[1] Alin C. Popescu, and Hany Farid, "Statistical Tools for Digital Forensics", Information Hiding, Springer Berlin Heidelberg, Vol. 3200, pp. 128-147, January 2005.

[2] Archana V Mir, S B Dhok, P D Porey, and N J Mistry, "Automation of JPEG ghost detection using graph based segmentation", International Journal of Computational Engineering Research, Vol. 3 No. 12, pp. 42-47, December. 2013.

[3] Diego Garcia-Ordas, Laura Fernandez-Robles, Enrique Alegre, Maria Teresa Garcia-Ordas, and Oscar Garcia-Olalla, "Automatic tampering detection in spliced images with different compression levels", Pattern Recognition and Image Analysis, Lecture Notes in Computer Science, Vol. 7887, pp. 416-423, 2013.

[4] Fabian Zach, Christian Riess, and Elli Angelopoulou, "Automated image forgery detection through classification of JPEG ghosts", Joint 34th DAGM and 36th OAGM Symposium Austria, Proceedings : Lecture Notes in Computer Science, Vol. 7476, pp. 185-194, August, 2012.

[5] Fangjun Huang, Jiwu Huang, and Yun Qing Shi, "Detecting Double JPEG Compression With the Same Quantization Matrix", IEEE Transactions on Information Forensics and Security, Vol. 5. No. 4, pp. 848-856, December 2010.

[6] Gerald Schaefer, and Michal Stich, "UCID–An uncompressed colour image database", Storage and Retrieval Methods and Applications for Multimedia, SPIE Vol. 5307, 2004.

[7] Hany Farid, "Exposing digital forgeries from JPEG ghosts", IEEE Transaction on Information Forensics and Security, Vol. 4, No. 1, pp. 154-160, March. 2009.

[8] Hany Farid, "Image forgery detection : A survey", IEEE Signal Processing Magazine, pp. 16-25, March 2009.

[9] Junfeng He, Zhouchen Lin, LifengWang, and Xiaoou Tang, "Detecting doctored JPEG images via DCT coefficient analysis", European Conference on Computer Vision, Austria, ECCV Proceedings, Part III, pp. 423–435, May 2006.

[10] Junfeng He, Zhouchen Lin, LifengWang, and Xiaoou Tang, "Identification of cut & paste tampering by means of double JPEG detection and image segmentation", IEEE International Symposium on Circuits and Systems, pp. 1687-1690, June 2010.

[11] Mandeep kaur, Jyoti, and Prakriti, "Image Tamper Detection based on JPEG Artifacts", International Journal of Application or Innovation in Engineering & Management, Vol. 3, No. 4, pp. 358-363, April 2014.

[12] Vrizlynn L. L. Thing, Yu Chen, and Carmen Cheh, "An improved double compression detection method for JPEG image forensics", IEEE International Symposium on Multimedia, pp. 290 - 297, December 2012.

[13] W. Luo, Z. Qu, J. Huang, and G. Qiu, "A novel method for detecting cropped and recompressed image blocks,", IEEE Conf. on Acoustics, Speech and Signal Processing, Vol. II, pp. 217–220, April 2007.

Book References

[14] John Miano – *Compressed Image File Formats: JPEG, PNG, GIF, XBM, BMP*; First Edition; Addison Wesley Longman, Inc. Massachusetts, 1999.