Efficient Implementation of Computer Forensic Techniques for Forgery Detection in JPEG

Jyoti Kumari[1], Dr. Sunil Patil[2], Abhinav Shukla[3]
M.Tech (CSE) Scholar, Department of CSE, RKDF University Bhopal, India[1]
Professor,Department of CSE, RKDF University Bhopal, India[2]
Head of Department, Department of CSE, RKDF University Bhopal, India[3]

Abstract :- Anti-forensic methods are actions whose goal is to prevent proper forensic investigation process. In the field of anti-forensics, a set of techniques are designed to fool current forensic methodologies. An image's JPEG compression history is used to provide evidence of image manipulation, it provides information about the camera used to generate the particular image and identify forged regions within an image. When an image is JPEG compressed it would affect the visual quality of the image and forensic detectors are able to identify the traces of the image that it should be JPEG compressed or not. This paper focuses an anti-forensic method which aims at removing from an image the footprints left by JPEG compression in both spatial domain and DCT domain. It can defeat the existing forensic detectors which aim to identify the traces of JPEG compression history. It provides a better tradeoff between the forensic undetectability and the visual image quality.

Keywords-Anti-Forensics,JPEG compression, image manipulation, DCT

1. INTRODUCTION

The widespread availability of photo editing software reduces the difficulty to make visually plausible fake images. As a result, a number of researchers have developed many computer-based forensic algorithms to detect digital forgeries. Additionally, several methods have been developed to perform other forensically significant tasks such as tracing an image's processing history or determining the device used to capture an image. Anti-forensic [1]operations may be used to provide intellectual property protection by preventing the reverse engineering of proprietary signal processing operations used by digital cameras through digital forensic means.

Many of the digital forensic techniques rely on detecting artifacts left in an image by JPEG compression[2]. The JPEG format is widely used as one of the most popular lossy

image compression formats today, and is adopted by various digital cameras and image processing software tools.

This paper proposes a method to remove from a given image the footprints left by JPEG compression in both spatial domain and DCT domain. With reasonable

loss of image quality, this anti forensic method can defeat existing the forensic detectors that attempt to identify traces of the image JPEG compression history. An image's JPEG compression history[6] which is used to provide

evidence of image manipulation, supply particular information about the camera used to capture an image and identify forged regions within an image. In this method firstly because of a total variation-based deblocking operation, partly recovered DCT information is then used to build an adaptive local dithering signal model, which would be able to bring the DCT histogram of the processed

image close to that of the original one. After that, a perceptual DCT histogram smoothing is carried out by solving a simplified assignment problem, where cost function is established as the total perceptual quality loss due to the DCT coefficient modification. Then the second-round deblocking and decalibration operations successfully bring the image statistics that are used by JPEG forensic detectors to the normal status. It provides a better tradeoff between the forensic undetectability and the visual image quality. Also this method is used in disguising double JPEG compression artifacts. This method can also reduce the computation cost.

2. RELATED WORK

Sometimes image processing inherits images so that processing is to be carried without knowledge of past operations. To carry out further processing it is useful to know whether the image has been JPEG compressed or not. Many of digital forensic techniques rely on detecting artifacts left in an image compression. Methods designed to detect previous instances of JPEG compression in images saved using uncompressed or losslessly compressed file formats. The JPEG format is widely used as one of the most popular lossy image compression format. Much of processing is focused that whether an image is JPEG compressed or not for forensic purposes.

In JPEG compression history, to fill the gaps in the comb like distribution of DCT coefficients, a dithering[3] operation is used. This operation is used to conduct DCT histogram smoothing based on Laplacian model. This dithering operation successfully fools the detector in DCT domain artifacts. But it leaves the footprints which can be detected by advanced detectors.

Another disadvantage is the degradation of the image visual quality. In the case of double JPEG compression artifacts may appear at least in one part of the resulting manipulated[7] image. For a forger JPEG anti forensics may be useful to hide traces of the first JPEG compression. Therefore double JPEG compression artifacts are less likely to appear as evidence of tampering.

3. PROPOSED WORK

The original uncompressed image is firstly split into L non-overlapping 8 × 8 pixel value blocks. For each block, a 2-dimensional DCT is afterwards applied to obtain its corresponding DCT coefficient block. As DCT is an orthogonal linear transform, this mapping can be modeled as a matrix multiplication. Then rounding function is performed. The resulting, quantized DCT coefficients are then losslessly encoded. As to the decompression, the quantized DCT coefficient is extracted from the decoded bit stream, and then dequantized by multiplying it by the corresponding quantization step. The dequantized DCT coefficients are then transformed to the spatial domain by the 2-D

inverse discrete cosine transform (IDCT), which can be modeled as multiplication by the 8×8 block IDCT matrix. At last rounding and truncation operation is applied to constrain the pixel values to be integers within [0, 255], and the decoded JPEG image is obtained as *J*.

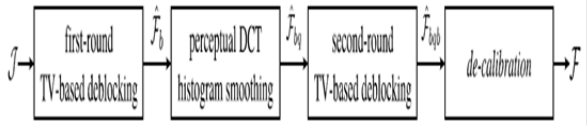
J
first-round TV-based deblocking
$\hat{F}_b$
perceptual DCT histogram smoothing
$\hat{F}_{bq}$
second-round TV-based deblocking
$\hat{F}_{bqb}$
de-calibration
F

Fig 1: JPEG Forgery Creation Process

An extended four-step procedure is proposed, which is composed of total variation (TV)-based deblocking, perceptual DCT histogram smoothing based on an adaptive local dithering signal model, second round TV- based deblocking and decalibration. The effectiveness of the proposed anti-forensic method is confirmed by its undetectability against existing JPEG forensic detectors in both the spatial and the DCT domains. The first step is TV-based deblocking in the spatial domain. Besides the removal of JPEG blocking artifacts, another purpose of this step is to partly and plausibly fill gaps in the DCT histogram, so as to facilitate the following step of explicit histogram smoothing. Experimentally, it is necessary and beneficial to conduct this first-round deblocking, especially for a better histogram restoration in the high frequency sub bands where all DCT coefficients are quantized to zero in the JPEG image.

In the deblocked image $\hat{F}_b$, the comb-like DCT quantization artifacts are no longer as obvious as those in the JPEG image *J*. Under the hypothesis that the partly recovered DCT-domain information is reliable, the next step naturally goes to further filling the remaining gaps in the DCT histogram. This leads us to the construction of an adaptive local model for the DCT coefficient distribution, with which a perceptual histogram mapping method is thereafter proposed to modify the DCT coefficients while minimizing the total SSIM (structural similarity) value loss.

The removal of the DCT quantization artifacts is at the cost of introducing a small amount of unnatural noise and blocking artifacts in the spatial domain to the output image $\hat{F}_{bq}$. Hence, move to the spatial domain again and conduct a second round TV-based deblocking and regularization. The resulting image $\hat{F}_{bqb}$ is at last processed by the decalibration operation to generate the JPEG forgery *F*.

4.ARCHITECTURAL DIAGRAM

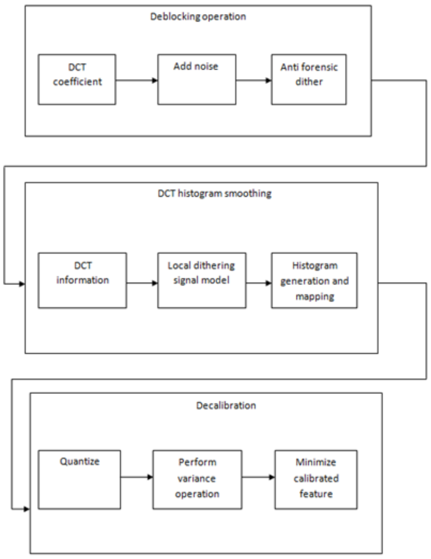

Fig 2. Architecture

In the deblocking operation if a previously JPEG compressed image is to be passed off as never having undergone compression, JPEG blocking artifacts must be removed from the image after anti-forensic dither has been applied to its DCT coefficients. There are many number of deblocking algorithms have been found since the introduction of the JPEG compression standard, but majority of these are ill suited for anti-forensic purposes. Considering for an anti-forensic deblocking operation to be successful, it must have to remove all visual and statistical traces of block artifacts without resulting in forensically detectable changes to an image's DCT coefficient histograms. By lightly smoothing an image followed by adding low-power white Gaussian noise, it would be able to remove statistical traces of JPEG blocking artifacts without causing the images DCT coefficient distribution to deviate from the Laplace distribution.

After JPEG image has been processed using the deblocking method, the gaps in the DCT domain have been partly filled in the obtained image. In order to achieve a better forensic undetectability, it is necessary to fill the gaps left in the DCT histogram of obtained image. The partly recovered information in the DCT domain will help us to build an adaptive local dithering signal model based on both the Laplacian distribution and the uniform distribution for a better goodness-of-fit. In DCT histogram mapping it generates the dithered signal using the adaptive local dithering signal model. The distribution of DCT coefficients of obtained image is minimized by the introduced distortion in the spatial domain, by solving an assignment problem whose cost function is defined as the total perceptual quality loss. After solving a simplified assignment problem it would be able to smooth the DCT histogram with minimum introduced distortion in the spatial domain.

The resulting image is at last processed by the decalibration operation. It is hard to further decrease this value by performing deblocking, while keeping good visual quality. The decalibration operation is done by optimizing the energy function and takes the variance measure. In order to fool the detector, a random threshold for each image is drawn from the distribution of the calibrated feature values for genuine, uncompressed images, and the iteration stops once the calibrated feature value drops below it. After decalibration, the JPEG forgery is obtained.

The proposed scheme is constructed in following four steps

1. Image preprocessing
2. DCT histogram generation
3. Quantization mechanism
4. Comparison of matrices

1. ***Image preprocessing***

Browse the concerned image. For an image, consider the grayscale image of each picture element. Split the original uncompressed image in pixels. Apply a DCT to blocks of pixels, thus removing redundant image data. The image data is divided up into 8*8 blocks of pixels. From this point on each of the color component is processed independently, so a pixel means a single value, even in the color image. A DCT is applied in each 8*8 blocks.DCT converts the spatial image representation into a frequency map, the lower order or DC term represents the average value in the block, While successful higher order(AC) terms represent the strength of more and more rapid changes across the width or height of block. The highest AC term represents the strength of a cosine wave alternating from maximum to minimum at adjacent pixels. The DCT calculation is fairly complex in fact this is the most costly step in JPEG compression. Discard high frequency data easily without losing low frequency information. The DCT step itself is a lossless except for round off errors. Thus corresponding DCT coefficient block can be obtained. Quantize each block of DCT coefficients. Perform one to one division and round off. Encode the resulting coefficients of image data.

2. ***DCT histogram generation***

For the generation of the histogram RGB color values of concerned image is considered. It is transformed into luminance. The color space transformation is performed on pixel by pixel basis. DCT coefficient calculated value can be plotted with the DCT coefficient frequency. The histogram of a gray-scale image consists of a discrete array of bins, each representing a certain gray- level range and storing the number of pixels in the image whose gray-level falls into that range. In other words, it defines a discrete function that maps a gray-level range to the frequency of occurrence in the image. Scatter-based histogram generation consists of two sub-tasks: Bin selection for each input pixel and accumulation of bin contents. It renders one point primitive for each input pixel. Then compute the bin index in the vertex shader and convert it to an output location that maps into our 1D bin texture. The fragment that is rasterized into our desired bin location in the histogram render target is accumulated by configuring the hardware blend units to add the incoming fragment to the contents of render target. After scattering and accumulating all points in this manner, the output render target will contain the desired histogram. By looking at the histogram for a specific image a viewer will be able to judge the entire tonal distribution at a glance.

3. ***Quantization mechanism***

In the JPEG compression standard, the input image is first divided into non overlapping pixel blocks of size 8*8. For each block, the two-dimensional discrete cosine transform (DCT) is computed. There is a one-to-one mapping between the index and the position of a coefficient within a DCT block. Each DCT coefficient is quantized with a quantization step size. In many JPEG implementations, it is customary to define as a scaled version of a template matrix, by adjusting a (scalar) quality factor.

For instance, the quantization matrices adopted by the Independent JPEG Group (IJG) which are obtained by properly scaling the image-independent quantization matrices of the JPEG standard. The quantization levels are obtained from the original coefficients. The quantization levels are entropy coded and written in the JPEG bit stream. When the bit stream is decoded, the DCT values are reconstructed from the quantization levels. Then, the inverse DCT is applied to each block, and the result is rounded and truncated in order to take integer values on [0,255].

Due to the quantization process, the dequantized coefficients can only assume values that are integer multiples of the quantization step size. The process of rounding and truncating the decompressed pixel values perturbs the comb-shaped distribution.

To conceal the traces of JPEG compression by filling the gaps in the comb-shaped distribution of by adding a dithering, noise-like, signal in such a way that the distribution of the dithered coefficients approximates the original distribution. The addition of the anti-forensic dither corresponds to injecting a noise-like signal in the pixel domain.

As a result, the dithered image is distorted with respect to the JPEG-compressed image. Then characterize analytically the distortion in the DCT domain, showing that it is a function of both the distribution of the original transform coefficients and the quantization step size. The energy of anti-forensic dithering is concentrated in the middle DCT frequencies, thus resulting in a grainy noise in the spatial domain. With this fact to select a proper set of DCT coefficients to analyze in order to detect JPEG anti-forensics.

Quantization table matrices are employed for luminance and chrominance data with chrominance data being quantized more heavily than the luminance data. It allows JPEG to exploit further the eye’s differing sensitivity to luminance and chrominance. Quantization mechanism that controls the quality setting of most JPEG compressors.

4. ***Comparison of matrices***

In this step with the comparison of matrices we can find the incorrectly classified images. It can be done with comparing of image properties, histogram values, DCT comparison, noise presence etc. Image properties can have the comparison with the dimension, extension of the images. In histogram comparison RGB values are compared, anti-forensically it would be difficult to find this comparison. Next DCT values are compared, the original DCT values and modified DCT values are compared and find the differences. Finally check the presence of noise, if the image is anti forensically treated there would be differences in the image matrices values, otherwise the image is original.

5.CONCLUSION

This paper mainly focuses on removing from a given image the footprints left by JPEG compression. JPEG compression is one of the most promising applications in the area of image processing. This anti forensic method can provide better visual quality to the concerned image. During comparison of original image and suspected image if noise is detected the image has been once compressed and anti forensically treated otherwise image is original. Interactive applications, for example in forensic security, also benefit from the advances in the traces of JPEG compression history.

ACKNOWLEDGEMENT

This work was supported in part by the Department of Computer Science & Engineering, of the University. I would like to show my gratitude to the guide and the HoD for their valuable guidance

REFERENCES

[1] W. Fan, K. Wang, F. Cayre, and Z. Xiong, "A variational approach to JPEG anti-forensics," in Proc. IEEE Int. Conf. Acoust., Speech, Signal Process., May 2013, pp. 3058–3062.

[2] Z. Fan and R. L. De Queiroz, "Identification of bitmap compression history: JPEG detection and quantizer estimation," IEEE Trans. ImageProcess., vol. 12, no. 2, pp. 230–235, Feb. 2003.

[3] G. Valenzise, M. Tagliasacchi, and S. Tubaro, "The cost of JPEG compression anti-forensics," in Proc. IEEE Int. Conf. Acoust., Speech, Signal Process., May 2011, pp. 1884–1887.

[4] C. Chen and Y. Q. Shi, "JPEG image steganalysis utilizing both intrablock and interblock correlations," in Proc. IEEE Int. Symp. Circuits Syst., May 2008, pp. 3029–3032.

[5] T. Bianchi and A. Piva, "Detection of nonaligned double JPEG compression based on integer periodicity maps," IEEE Trans. Inf. Forensics Security, vol. 7, no. 2, pp. 842–848, Apr. 2012

[6] M. Stamm, S. Tjoa, W. S. Lin, and K. J. R. Liu, "Anti-forensics of JPEG compression," in Proc. IEEE Int. Conf. Acoust., Speech, Signal Process., Mar. 2010, pp. 1694–1697.

[7] 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.

[8] M. Stamm, S. Tjoa, W. S. Lin, and K. J. R. Liu, "Undetectable image tampering through JPEG compression anti-forensics," in Proc. 17th IEEE Int. Conf. Image Process., Sep. 2010, pp.2109–2112.