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

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 2, Issue 3, March 2013

Forensic Technique for Detecting Tamper in Digital Image Compression

Abhitha.E[1] V.J Arul Karthick[2]

PG-Scholar, Electronics and communication, SNS College of technology, Coimbatore, India[1]

Assistant professor, Electronics and communication, SNS College of technology, Coimbatore, India[2]

Abstract: As society has become increasingly depend upon digital images to communicate visual information. Nowadays one of the principal means for communication is digital visual media. In most digital image communication the main problem is its authenticity. For the verification of image integrity [4], authentication, and tampering [3] detection a number of forensic techniques have been developed. Digital image forensics [7] is a brand new research field which aims at finding the authenticity of images by recovering information about their history. Digital Image Forensics which is that branch of multimedia security, combined with Digital Watermarking, gives at contrasting and effective image manipulation .Most of the image manipulations occurs at the time of compression. That is Image manipulations means changing any of the DCT and DWT coefficients. A set of forensic techniques can be developed by analysing these coefficients. This work can be implemented over JPEG, JPEG2000, SPHIT, EZW compression. In this paper we discussed forensic techniques in SPHIT image compression.

Keywords: Digital forensics, image compression, SPIHT compression, DWT coefficients.

I. INTRODUCTION

In recent years due to the widespread availability of digital cameras and the development of multimedia yields the usage of digital image communication. But number of photo editing and manipulations are available. This will easily affect the image, where the authenticity of digital images is often in doubt. These manipulations rems as image forgery or tampering of an image. To prevent the alteration of images researchers have developed a variety of digital image forensic techniques [1]. These techniques are designed to identify an image's originating camera, trace its processing history and determine its authenticity. Each and every image has its own unique property. This unique property defined as fingerprints of an image. Each image has its own fingerprint. Commonly the fingerprints are divided into extrinsic fingerprints and intrinsic fingerprints [2]. Extrinsic fingerprints means the digital signature, and some watermarking etc. But intrinsic finger prints mean the coefficients present in that image. In this paper we discussed about the intrinsic fingerprint operation. The image editors can able change the intrinsic fingerprints at the time of compression. So this paper is developed for detecting any image tampering occurs at the time of compression. Here this work is implemented over SPIHT(set partitioning in hierarchical tress).

In a SPHIT compressed images the intrinsic fingerprint is DWT confidents. Any change in the DWT coefficients of compressed and uncompressed images can be detected from estimating the histogram. This paper can be implemented using MATLAB software.

II. DIGITAL IMAGE FORENSICS

Images and videos have become the main digital information carriers in the digital word. The high potential of visual media and the ease in their capture, distribution and storage is such that they are utilized to convey information. Image processing experts can very easily access image content or modify the images, without leaving visually detectable traces. Moreover, in recent years there is a spread of low-cost, user friendly editing tools, the tampering arts and counterfeiting visual content is not restricted to image experts. As an effect, the modification of images for malicious purposes is nowadays more common .Digital image forensics (DIF) providing a tools to support blind investigation [7] .Two principal methods evolve under Digital Image Forensics. The first one includes methods to identify the image capturing device, or determine which devices did not capture it. These methods called as image source device identification techniques. The second group of methods aims to studying inconsistencies in natural image statistics. Commonly we will refer to these methods as tampering detection techniques.

The insertion in an image of material originally coming from another source is one of the powerful tools to subvert the message contained in visual media. Modern techniques and image editing software allow changing of image composites obtaining results that are hardly detectable by the human eye .Blending and matting[7] techniques are applicable to mask the boundaries of the combined regions and to give the image a more uniform aspect. Also, the creation of image components require geometric transformation. Rotation, translation and scaling are often needed to make sure that the spliced object respects the original image view and scale. Geometric transforms involve re-sampling, also called interpolation (e.g. nearest neighbour, bilinear, bicubic). The re-sampling process generates artifacts in the image histogram, and hence provides a very useful key for compositing detection.

III. WAVELET DECOMPOSITION OF IMAGES

Wavelets are mathematical functions that decompose data or image into different frequency bands or components, and then study each component with a resolution matched to its scale. Wavelets have advantages over Fourier transform, wavelet applicable in where the signal contains discontinuities and sharp spikes. In past wavelets are used for in the fields of mathematics, physics and electrical and instrumentation engineering. But now the wavelet transform have new application the field of digital image processing, turbulence, human vision, radar, and some natural calamities prediction.

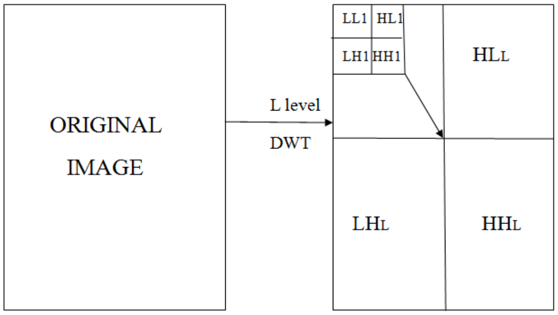


Fig.1 Decomposition of image

The wavelet transformation is a mathematical tool for decomposition of an image. The wavelet transform is a hierarchical system identical to sub band filtering system, in which sub bands are logarithmically spaced in frequency domain. The basic idea of the DWT for a two-dimensional image is explained as follows. An image is first decomposed into four parts based on frequency sub bands, by sub sampling horizontal and vertical components using sub band filters and named as Low-Low (LL), Low-High (LH), High- Low (HL), and High- High (HH) sub bands as shown in figure 1. To obtain the next set of scaled wavelet coefficients, second level decompositions are needed. In second level decomposition the first sub band LL is further decomposed and critically sub sampled. This process is repeated several times in order to get different sub bands. The block diagram of this image decomposition is shown in figure 1. Each level has various sub bands information such as low– low, low–high, high–low, and high–high frequency bands. From these DWT coefficients, the original image can be reconstructed. This process is called the inverse DWT (IDWT).

IV. SET PARTITIONING IN HIERARCHICAL TREES

Set partitioning in hierarchical tress is a powerful image compression method based on wavelet transform. The principles are partial ordering of coefficients by magnitude using a set partitioning sorting algorithm, ordered bit plane transmission, and advantages of self-similarity across different scales of an image wavelet transform. The SPIHT is a traditional methods for image compression and it is more advance in the field of compression. SPIHT gives higher The SPIHT process represents a very effective form of entropy-coding. Mainly two forms of coding: binary-encoded (extremely simple) and context-based adaptive arithmetic coded (sophisticated).SPHIT has greater coding/decoding speed it results compression simplicity. In SPIHT time to encode is nearly equal to the time to decode, so the algorithm called symmetric algorithm. The SPIHT technique is mainly based on three concepts:

1) Partial ordering of the transformed image elements by its magnitude, with transmission of order using subset partitioning algorithm that is duplicated at the decoder,
2) Bit plane ordered transmission of refinement bits.
3)Exploitation of the self-similarity of the image wavelet transform across different scales.

A. *Progressive image transmission*

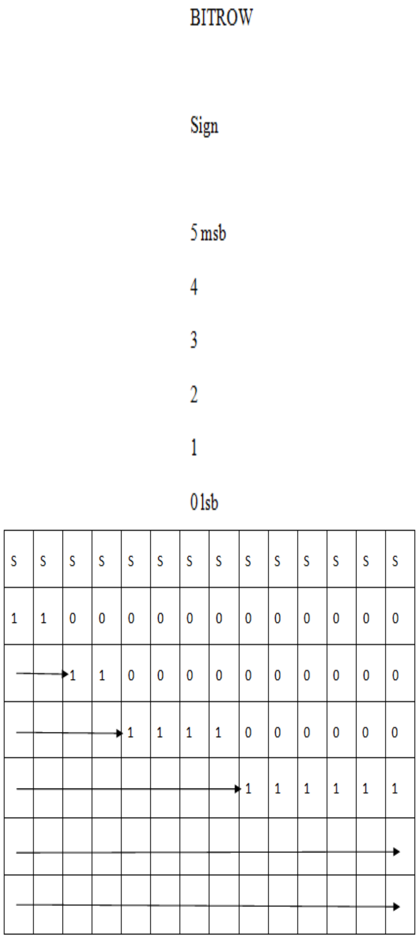
The coding is done to the array

$$\mathbf{c} = \Omega(\mathbf{p}) \qquad (1)$$

where $\Omega(\cdot)$ represents a unitary hierarchical sub band transformation and **p** represents the original image.**'d'** represents the reconstruction vector.In a progressive transmission scheme, the decoder initially **d** to zero and updates its components according to the coded message. Reconstructed image can be obtain after receiving the value (approximate or exact) of some coefficients,

$$\mathbf{r} = \Omega^{-1}(\mathbf{d}) \qquad (2)$$

A major objective in a progressive transmission scheme is to select the most important information. This indicates that the coefficients with larger magnitude should be transmitted first because they have a larger amount of information. The information in the value of **c** can ranked according to its binary representation, and the MSB bits should be transmitted first then LSB. Fig. 2 indicates the schematic binary representation of a list of magnitude-ordered coefficients



s	s	s	s	s	s	s	s	s	s	s	s	s	s
1	1	0	0	0	0	0	0	0	0	0	0	0	0
→	→	1	1	0	0	0	0	0	0	0	0	0	0
→	→	→	→	1	1	1	1	0	0	0	0	0	0
→	→	→	→	→	→	→	→	1	1	1	1	1	1
→	→	→	→	→	→	→	→	→	→	→	→	→	→
→	→	→	→	→	→	→	→	→	→	→	→	→	→

Fig. 2 Binary representation of a list of magnitude-ordered coefficients

B. *Set partitioning sorting algorithm*

One of the main features of the proposed coding method is that the ordering data is not transmitted. The transmission based on the fact that the execution path of any algorithm is defined by the results of the comparisons on its branching points. So, same sorting algorithm used for encoder and decoder .An algorithm should require that simply selects the coefficients such that $2^n \leqq |c_{i,j}| \leqq 2^{n+1}$,E. with n decremented in each pass. Given n, if $|c_{i,j}| \geqq 2^n$ then the coefficient is significant; otherwise it is called insignificant. A set of pixels divided into partitioning subsets τ_m according to sorting algorithm and performs the magnitude test.

$$\max_{(i,j)\in\tau_m}\left\{\left|c_{i,j}\right|\right\}\geq 2^n \,? \qquad (3)$$

If the decoder receives a 'no' then it knows that all coefficients in τ_m are insignificant. If the answer is 'yes' then the subset is significant, then a certain rule applied by the encoder and the decoder is used to partition the new subsets. This set partition process continues til the magnitude test is done to all single coordinate significant subsets in order to identify each significant coefficient is a subset .To decrease the number of magnitude comparisons(message bits) we define a set partitioning rule based on sub band pyramid. The aim is to create new partitions such that subsets expected to be large number insignificant of elements, and subsets expected to be significant contain only one element. To make a clear relationship between magnitude comparisons and message bits.

C. *Spatial orientation of trees*

Normally, low frequency components contains most of an image's energy is concentrated. Consequently, as we move

from the highest to the lowest levels of the sub band pyramid variance reduced. A tree structure, called spatial orientation tree (SOT), defines the spatial relationship on the hierarchical sub band pyramid.The new coding method contained following parameters:
$O(i, j)$: set of coordinates of all offspring of node (i, j);
$D(i, j)$: set of coordinates of all descendants of the node (i, j);
H : set of coordinates of all spatial orientation tree roots (nodes in the highest pyramid level);

$$L(i, j) = D(i, j) - O(i, j). \qquad (4)$$

The set partitioning rules are simply the following:

- The initial partition is formed with the sets $\{(i, j)\}$ and $D(i, j)$, for all $(i, j) \in H$.
- If $D(i, j)$ is significant, then it is partitioned into $L(i, j)$ plus the four single-element sets with $(k, l) \in O(i, j)$.
- If $L(i, j)$ is significant, then it is partitioned into the four sets $D(i, j)$, with $(k, l) \in O(i, j)$.

D. *Coding algorithm*

The coefficients are stored in three ordered lists, called list of insignificant sets (LIS), list of insignificant pixels (LIP), and list of significant pixels (LSP). In all lists each entry is identified by a coordinate (i, j), which in the LIP and LSP gives individual pixels, and in the LIS represents either the set $D(i, j)$ or $L(i, j)$. To differentiate between them, LIS entry is of type A if it represents $D(i, j)$, and of type B if it represents $L(i, j)$. During the sorting pass, those that become significant are moved to the LSP. Similarly, sets are sequentially evaluated according to LIS order, and when a set is found to be significant it is removed from the list and partitioned. More than one element in the new subset are added back to the LIS, mean while the single-coordinate sets are added to the end of the LIP or the LSP, whether they are insignificant or significant, respectively. The LSP consist of the coordinates of the pixels that are visited in the refinement pass.

V. Forensic detection

Image forgery detection techniques which developed by detecting the intrinsic fingerprint of each operation. Image operations such as histogram calculation, contrast enhancement and quantization table matching. Here we use the histogram estimation. For forensic techniques DWT coefficients are used. Histogram of DWT coefficients gives intrinsic fingerprint of an image. Coefficients value verses number of coefficients are plotted for both compressed version and uncompressed version of images. Finally we calculate the number of coefficients values for original version of compressed images. Save this value in a workspace. And then calculate the number of coefficients values for tampered compressed images. After this step calculate the difference between number of coefficients between tampered and original version, Estimate this value against a threshold value and detect the image manipulations.

VI. . Anti forensics

Anti-forensic is capable of fooling the our exiting forensic techniques. Anti-forensic operations designed to hide fingerprints of image manipulation may be applied to an image. The anti-forensic detection algorithm is purely based of probability distribution of coefficients. Adding some amount of noise called anti forensic dither to the compressed image so that distribution strictly matches to the estimated one. The distribution of original image and estimated image are same, so the forensic techniques failed to determine the modification. Anti forensics operation leaves its own compression fingerprints. A new techniques have been proposed to determine the effect of anti forensics. Anti-forensic image processing operations must be developed and studied so that weaknesses in existing image forensic techniques can be made known to researchers. This will allow researchers to know when forensic results can be trusted and to assist researchers in the development of improved digital forensic techniques. The study of anti-forensic operations may also lead to the development of techniques capable of detecting when an anti-forensic operation has been used.

VII. Simulation study and results

Experiments are performed on gray level images to verify the proposed method. These images are represented by 8 bits/pixel and size is 512 x 512. Image used for experiments are shown in below figure.

Fig. 3 Original image

If this image undergoes wavelet transform and then SPHIT coding algorithm,the number of bits/pixel has been decreased.Thus the compressed version of image contains reduced bits/pixel.For reconstruction of this compressed image first we apply the decoding and then reverse DWT

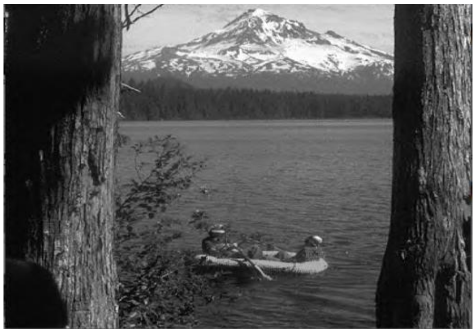


Fig.4 Reconstructed version of image

If the image editor's remove any of the compression finger prints from an image during the time of compression, this is called image tampering or forgery. The reconstructed version of image after removing the any one of the fingerprints shown in below. In some cases, depending on the type of edition the appearance of image will be varying.

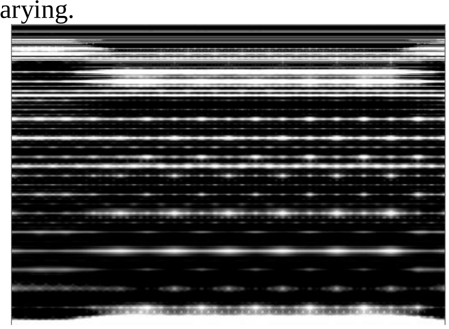


Fig.5 Reconstructed versions after removing Compression fingerprints

Histogram of coefficient values from the DWT taken from an uncompressed version of the image shown below

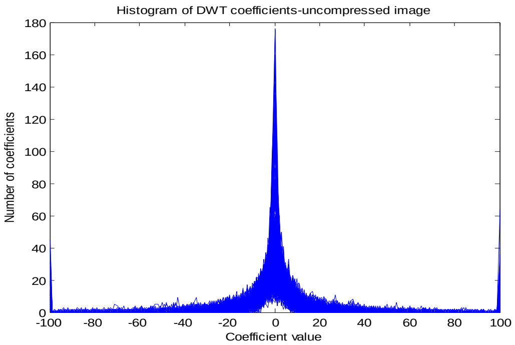


Fig.6 Histogram of uncompressed image

To analyse the DWT coefficients after compression we have to plot the histogram of compressed version of an image. Histogram of coefficient values from the DWT taken from a compressed version of the image shown below

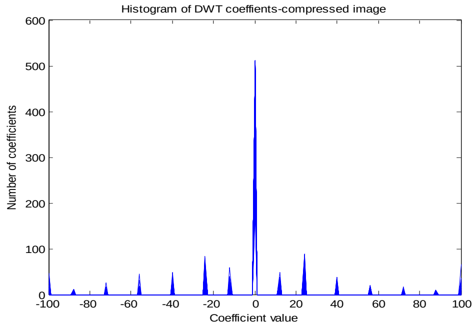


Fig.7 Histogram of compressed image

Histogram of coefficient from the DWT sub band taken from an compressed version of the image after removal of compression fingerprint is shown below

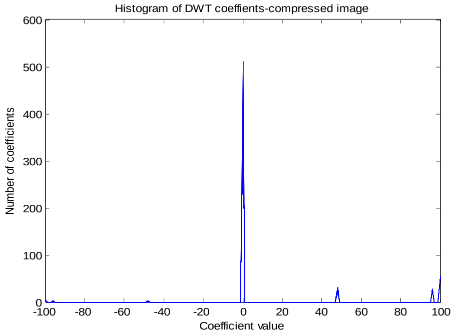


Fig.8 Histogram of tampered compressed image

By comparing the figures 8 & 9 we can detect the image tampering or manipulations before decompression.

Anti forensically modified image after the addition of anti forensic dither into the original wavelet coefficients are show below. This image is looks likes the original one, but actually it is the tampered version of image. This fools the existing forensic operation.

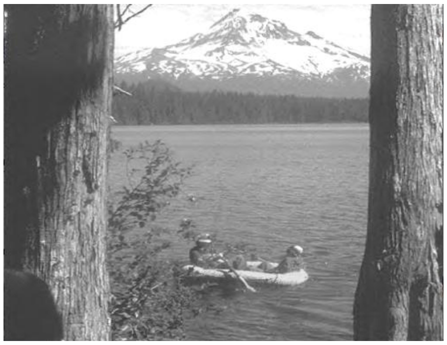


Fig.9 Anti forensically modified image

To analyse the three types of image that is original image, tampered version and anti forensically modified image ,here considered the PSNR(Peake Signal to Noise Ratio) and MSE(Mean Square Error

TABLE I:
PSNR AND MSE ANALYSIS

Different types of images	**MSE**	**PSNR**
Original image	31.4363	33.1565
After image editing	4.0016e+003	12.1084
Anti forensically modified image	31.6650	33.1250

VIII. Conclusion

The algorithm that operates through set partitioning in hierarchical trees (SPIHT) which is a efficient method of compression. In this paper, we have proposed a set of forensic operations capable of finding compression fingerprints from digital images. To do this, we developed a generalized framework for the removal of intrinsic fingerprints from an image's transform coefficients. In future work my work is extending in to detecting anti-forensic dither. In this developing a new algorithm that identifies the anti-forensic effect .Anti-forensic is capable of fooling forensic techniques. In proposed system techniques have been developed to detect the use of anti-forensics. An anti-forensic operation leaves behind its own unique fingerprints, a new forensic detection technique can be designed by using this fingerprints

ACKNOWLEDGMENT

The authors would like to thank to SNS groups of institutions and Anna University for their support to our work.

REFERENCES

[1] Anti-Forensics of Digital Image Compression Matthew C. Stamm, Student Member, IEEE, and K. J. Ray Liu, Fellow, IEEE Trans. Inf. Forensics and Security, vol. 6, no. 3, Sep 2011

[2] A. Swaminathan,M.Wu, asnd K. J. R. Liu, "Digital image forensics via intrinsic fingerprints," IEEE Trans. Inf. Forensics Security, vol. 3, no. 1, pp. 101–117, Mar. 2008.

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

[4] M. Chen, J. Fridrich, M. Goljan, and J. Lukáˇs, "Determining image origin and integrity using sensor noise," IEEE Trans. Inf. Forensics Security, vol. 3, no. 1, pp. 74–90, Mar. 2008.

[5] M. C. Stamm, S.K. Tjoa,W. S. Lin, and K. J.1] R. Liu, "Anti-forensics of JPEG compression," in Proc. IEEE Int. Conf. Acoust., Speech, Signal Process., Mar. 2010, pp. 1694–1697.

[6] W. S. Lin, S. K. Tjoa, H. V. Zhao, and K. J. R. Liu, "Digital image source coder forensics via intrinsic fingerprints," IEEE Trans. Inf. Forensics Security, vol. 4, no. 3, pp. 460–475, Sep. 2009.

[7] Judith A. Redi & Wiem Taktak & Jean-Luc Dugelay, Digital image forensics: a booklet for beginners

[8] A. Skodras, C. Christopoulos, and T. Ebrahimi, "The JPEG 2000 still image compression standard," IEEE Signal Process. Mag., vol. 18, no.5, pp. 36–58, Sep. 2001..

[9] M. C. Stamm and K. J. R. Liu, "Wavelet-based image compression anti-forensics," in Proc. IEEE Int. Conf. Image Process., Sept. 2010, pp. 1737–1740.

[10] J. Lukáˇs and J. Fridrich, "Estimation of primary quantization matrix in double compressed JPEG images," in Proc. Digital Forensic Research Workshop, Aug. 2003, pp. 5–8.

BIOGRAPHY

Abhitha. E pursuing final year Master of engineering in Communication Systems, SNS College of Technology, Coimbatore.

V. J Arul Karthick working as Assistant Professor in SNS College of Technology, Coimbatore