Survey on Pixel and Format Based Image Forgery Detection Techniques

Anil Dada Warbhe
Asst. Professor, DEE,
MIET, Gondia

R. V. Dharaskar
Director, MPGI Integrated Campus, Nanded

ABSTRACT

In recent years, digital forensics emerged as a powerful and promising discipline to identify, detect and authenticate the digital images. This could be the authentic ground to present a proof of tempering as evidence in the court of law. The trust we have had till now in believing what we see started eroding. This is all happening due to the availability of the low cost, sophisticated yet easy to use tools and techniques. Due to the availability of these tools tempering the digital photographs getting easier and easier but at the same time it's very difficult to detect traces, if viewed by necked eye. Image forensic tools are mainly classified based on the approach used; active or passive. We here present a survey on pixel-based and format-based techniques, which comes under the realm of passive approach for digital image forgery detection.

General Terms

Digital Image Forensics, Image Processing, Forgery Detection Techniques.

Keywords

Digital image forensics, Image processing, Image tempering.

1. INTRODUCTION

We are living in the age where we are exposed every now and then to a variety of incredible digital images which are very hard to believe. Apparently, traditional saying, "seeing is believing", do not hold true. In everyday live starting from the magazines, to the mainstream media outlets, scientific journals, political campaigns, courtrooms and the photo hoaxes lands in our email inboxes, happening to be common and with increased frequency. Prior to the digital revolution, in the old times, it was very difficult to manipulate photographs taken by traditional film cameras due to the requirement of professional knowledge and sophisticated dark-room equipment.

With the availability of low cost off the shelf image manipulation and cloning tools, it is very easy to tamper and create fake images even to a person with lukewarm skills of photography. Every now and then we are been presented with the amazing and sometimes unbelievable kind of images in our email inboxes. Maximum of it are nothing but artificially synchronized photographic fakes, adopted, for promoting and floating different stories through media, emails and social networking websites.

The manipulations done in the images cannot be detected and make out by naked eyes; as manipulations may not leave obvious evidence of tampering. The manipulation to change the original content of the image is also known as image fakery. Image fakery is a cybercrime, but because of the lack of proper regulatory framework and infrastructure for prosecution of such evolving cybercrime, leads to dissatisfaction about increasing use of such tools. This scene developed the feeling of cynicism and mistrust among civilians.

Over the past few years, the field of digital forensics has emerged to help restore some trust to digital images. Hence, to detect such modifications in the original content in the images needs to be detected, and hence, the necessity of algorithms for efficiently verifying the integrity of images cannot be overemphasized in this digital era.

There are many methods or techniques for detecting tampered or forged image. Broadly, these methods can be classified into two major groups; Active Method and Passive Method. Active Method requires that certain information is embedded inside an image during the creation or before the image is being disseminated to the public. The information can be used to either detect the source of an image or to detect possible modification of an image. One of the techniques under active method is watermarking and other is a digital signature. On the other hand passive method do not necessarily need to have image with a digital watermark or digital signature, and having only the forged image in hand, it is possible to identify whether the given image is an authentic or forged one [1].

In the following sections we will discuss the general principles of the digital watermarking insertion and extraction, later the different techniques under passive methods.

2. WATERMARK BASED IMAGE AUTHENTICATION

Recently, numerous techniques for image integrity verifications have been proposed. Some techniques employ watermarking schemes to authenticate an image as well as determine its integrity. In this section we will briefly discuss the watermark based digital image authentication.

2.1 Watermark Insertion and Extraction

Watermark insertion at the source side include the generation of the watermark signal *W* and embed W in the original image *I* to get a watermarked image *I'*. The other side is to extract the watermark *W*, and give the confidence measure for the detected image [2]. Figure 1 shows the generic watermark embedding at the source side. We have the watermarked image $I' = f_1(I, W, K)$, where *K* denotes the key.

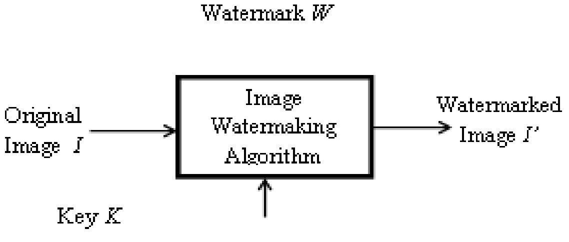

Fig 1: Generic watermark insertion

Once the watermark in embedded in the source image, if image undergoes tempering can be easily identified. The image I' can be easily authenticated by extracting the watermark embedded in it.

The watermark can be recovered as W'= f2 (I',K), where I' is the image to be authenticated. Watermark extraction is shown in Figure 2.

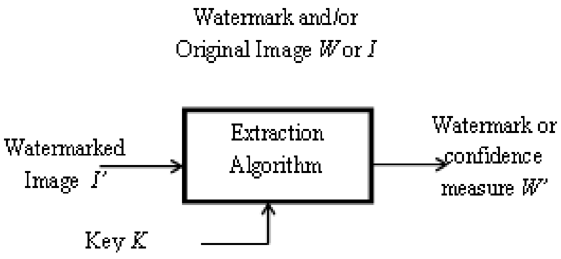

Fig 2: Watermark extraction

As mentioned above, the basic idea of the watermark based technology for image authentication is to add a watermark to the original image at the source side, and to recover the watermark fully or partly at the receiving side to identify whether the image has been altered. Therefore any manipulations before the watermark was embedded cannot be detected using this method. The signature-based method has a similar scheme and characteristics, and both of them are active methods [3].

3. PASSIVE FORENSIC TECHNIQUES

The drawback of the active method approach, especially in the watermark embedding is that a watermark must be inserted at the time of recording, which would limit this approach to specially equipped digital cameras. In contrast to these approaches, passive techniques for image forensics operate in the absence of any watermark or signature. These techniques work on the assumption that although digital forgeries may leave no visual clues that indicate tampering, they may alter the underlying statistics of an image. The set of image forensic uses techniques based on pixel- that detect statistical anomalies introduced at the pixel level; based on format that leverage the statistical correlations introduced by a specific lossy compression scheme; based on camera; based on principle of light source (Physics) and geometric-based techniques that make measurements of objects in the world and their positions relative to the camera [4].

In next section we will emphasize our review mainly two passive forensic technique based on pixel and format of the digital images.

4. PIXEL-BASED TECHNIQUES

When we are processing the images in the digital domain, the emphasis is on the pixel—the underlying building block of a digital image. In pixel based technique, here we describe four techniques for detecting various forms of tampering, each of which directly or indirectly analyzes pixel-level correlations that arise from a specific form of tampering.

4.1 Cloning

Cloning is, perhaps one of the most common image manipulations. The idea is to clone (copy and paste) portions of the image to conceal a person or object in the scene. When this is done with care, it can be difficult to detect cloning visually. And since the cloned regions can be of any shape and location, it is computationally impossible to search all possible image locations and sizes. Two computationally efficient algorithms have been developed to detect cloned image regions ([5], [6] ; see also [7], [8], and [9]). The authors in [5] first apply a block discrete cosine transform (DCT). Duplicated regions are detected by lexicographically sorting the DCT block coefficients and grouping similar blocks with the same spatial offset in the image. In a related approach, the authors in [6] apply a principal component analysis (PCA) on small fixed size image blocks to yield a reduced-dimension representation. Duplicated regions are again detected by lexicographically sorting and grouping all of the image blocks. Both the DCT and PCA representations are employed to reduce computational complexity and to ensure that the clone detection is robust to minor variations in the image due to additive noise or lossy compression.

4.2 Resampling

It is often necessary to resize, rotate, or stretch portions of an image in order to create a convincing composite. For example, when creating a composite of two people, one person may have to be resized to match the relative heights. This process requires resampling the original image onto a new sampling lattice, introducing specific periodic correlations between neighboring pixels. Because these correlations are unlikely to occur naturally, their presence can be used to detect this specific manipulation ([10]; related approaches are described in [11], [12], [13], and [14]).

A large range of re-samplings introduces some periodic correlations. If the specific form of the resampling correlations is known, then it would be straightforward to determine which pixels are correlated with their neighbors. If it is known which pixels are correlated with their neighbors, then the specific form of the correlations can easily be determined. But in practice neither is known. The expectation/maximization (EM) algorithm is used to simultaneously solve each of these problems. The EM algorithm is a two-step iterative algorithm: 1) in the expectation step, the probability of each pixel being correlated with its neighbor is estimated; and 2) in the maximization step, the specific form of the correlation between pixels is estimated. Assuming a linear interpolation model, the expectation step reduces to a Bayesian estimator, and the maximization step reduces to weighted least-squares estimation. The estimated probability is then used to determine if a portion of the image has been resampled.

4.3 Photomontage

A common form of photographic manipulation is the photomontage. Photomontage is a paste-up produced by sticking together photographic images, possibly followed by post- processing (e.g. edge softening and adding noise). It uses digital splicing of two or more images into a single composite. When performed carefully, the border between the spliced regions can be visually imperceptible. In [15] and [16], however, the authors show that splicing disrupts higher-order Fourier statistics, which can subsequently be used to detect splicing. Consider a signal $x(t)$ and its Fourier transform $X(w)$. The power spectrum $P(w)= X(w)\ X^*(w)$ is routinely used to analyze the frequency composition of a signal (* denotes complex conjugate). Moving beyond the power spectrum, the bispectrum can be expressed as

$$B(w_1, w_2)= X(w_1)\, X(w_2)\, X^*(w_1+ w_2) \qquad (1)$$

measures higher-order correlations between triples of frequencies w_1, w_2 and w_1+ w_2. Subtle discontinuities that result from splicing manifest themselves with an increase in the magnitude of the bispectrum and in a bias in the bispectrum phase, which are used to detect splicing in audio [15] and in images [16].

4.4 Statistical

There are a total of 256^{n^2} possible 8-bit gray-scale images of size n × n. With as few as n=10 pixels, there are a whopping 10240 possible images. If we were to draw randomly from this enormous space of possible images, it would be exceedingly unlikely to obtain a perceptually meaningful image. These observations suggest that photographs contain specific statistical properties.

The authors in [17], [18], and [19] exploit statistical regularities in natural images to detect various types of image manipulation. The authors in [17] compute first- and higher-order statistics from wavelet decomposition. This decomposition splits the frequency space into multiple scale and orientation subbands. The statistical model is composed of the first four statistical moments of each wavelet subband and higher-order statistics that capture the correlations between the various subbands. Supervised pattern classification is employed to classify images based on these statistical features. In a complementary approach, the authors in [18] construct a statistical model based on local co-occurrence statistics from image bit-planes. Specifically, the first four statistical moments are computed from the frequency of bit agreements and disagreements across bit planes. Nine features embodying binary string similarity are extracted from these measurements. Another eight features are extracted from the histograms of these measurements. The sequential floating forward search algorithm is used to select the most descriptive features, which are then used in a linear regression classifier for discriminating authentic from manipulated images. In both cases, the statistical model is used to detect everything from basic image manipulations such as resizing and filtering [18] to discriminating photographic from computer- generated images [20] and detecting hidden messages [21].

5. FORMAT BASED TECHNIQUES

JPEG is the most common format form images on the internet. JPEG is a lossy compression format; once we save the image any details that are lost cannot be recovered. In this view, lossy image compression schemes such as JPEG might be considered a forensic analyst's worst enemy. It is ironic, therefore, that the unique properties of lossy compression can be exploited for forensic analysis. Here we describe three forensic techniques that detect tampering in compressed images, each of which explicitly leverages details of the JPEG lossy compression scheme.

5.1 JPEG Quantization

Most cameras encode images in the JPEG format. This lossy compression scheme allows for some flexibility in how much compression is achieved. Manufacturers typically configure their devices differently in order to balance compression and quality to their own needs and tastes. As described in [22] and [23], this difference can be used to identify the source (camera make/model) of an image. Given a three-channel colour image (RGB), the standard JPEG compression scheme proceeds as follows: The RGB image is first converted into luminance/chrominance space (YCbCr). The two chrominance channels (CbCr) are typically subsampled by a factor of two relative to the luminance channel (Y). Each channel is then partitioned into 8 × 8 pixel blocks. These values are converted from unsigned to signed integers (e.g., from [0, 255] to [-128, 127]). Each block is converted to frequency space using a 2-D discrete cosine transform (DCT). Depending on the specific frequency and channel, each DCT coefficient, c, is then quantized by an amount q : |c/q|. This stage is the primary source of compression. The full quantization is specified as a table of 192 values—a set of 8 × 8 values associated with each frequency, for each of three channels (YCbCr). For low compression rates, these values tend toward a value of 1 and increase for higher compression rates. With some variations, the above sequence of steps is employed by JPEG encoders in digital cameras and photo-editing software. The primary source of variation in these encoders is the choice of quantization table. As such, a signature of sorts is embedded within each JPEG image. The quantization tables can be extracted from the encoded JPEG image or blindly estimated from the image, as described in [24]. The quantization tables can vary from within a single camera as a function of the quality setting, and while the tables are somewhat distinct, there is some overlap across cameras of different makes and models. Nevertheless, this simple observation allows for a crude form of digital image ballistics, whereby the source of an image can be confirmed or denied.

5.2 JPEG Blocking

As described in the previous sections, the basis for JPEG compression is the block DCT transform. Because each 8 × 8 pixel image block is individually transformed and quantized, artifacts appear at the border of neighbouring blocks in the form of horizontal and vertical edges. When an image is manipulated, these blocking artifacts may be disturbed. In [25], the authors characterize the blocking artifacts using pixel value differences within and across block boundaries. These differences tend to be smaller within blocks than across blocks. When an image is cropped and recompressed, a new set of blocking artifacts may be introduced that do not necessarily align with the original boundaries. Within- and across-block pixel value differences are computed from 4-pixel neighborhoods that are spatially offset from each other by a fixed amount, where one neighborhood lies entirely within a JPEG block and the other borders or overlaps a JPEG block. A histogram of these differences is computed from all 8 × 8 nonoverlapping image blocks. A 8 × 8 "blocking artifact" matrix (BAM) is computed as the average difference between these histograms. For uncompressed images, this matrix is random, while for a compressed image, this matrix has a specific pattern. When an image is cropped and recompressed, this pattern is disrupted. Supervised pattern classification is employed to discriminate between authentic and inauthentic BAMs.

5.3 Double JPEG

At a minimum, any digital manipulation requires that an image be loaded into a photo-editing software program and resaved. Since most images are stored in the JPEG format, it is likely that both the original and manipulated images are stored in this format. In this scenario, the manipulated image is compressed twice. Because of the lossy nature of the JPEG image format, this double compression introduces specific artifacts not present in singly compressed images (assuming that the image was not also cropped prior to the second compression). The presence of these artifacts can, therefore, be used as evidence of some manipulation [26], [27]. Note that double JPEG compression does not necessarily prove malicious tampering. For example, it is possible to inadvertently save an image after simply viewing it. As described before, quantization of the DCT coefficients c is the primary manner in which compression is achieved, denoted as, $q_a(c)=|c/a|$, where a is the quantization step (a strictly positive integer). Dequantization brings the quantized values back to their original range: $q^{-1}{}_a(c)=ac$. Note that quantization is not invertible, and that dequantization is not the inverse function of quantization. Double quantization that results from double compression is given by: $q_{ab}(c)=||c/b|b/a|$, where a and b are the quantization steps. Double quantization can be represented as a sequence of three steps: 1) quantization with step b, followed by 2) dequantization with step b, followed by 3) quantization with step a. In double quantization, the periodicity of the artifacts gets introduced into the histograms. It is this periodicity that the authors in [27] exploited to detect double JPEG compression. The work of [28] extended this approach to detect localized traces of double compression.

6. CONCLUSION

There is a growing need for digital image tampering detection techniques. Some of techniques, which were introduced in this paper, have been proposed to address various aspects of digital image tampering detection. From this survey, we can find that most proposed tampering detection methods aim at detecting inconsistencies in an image, and the majority of them belong to the low level category. Although many of these techniques are very promising and innovative, they have limitations and none of them by itself offers a definitive solution.

7. REFERENCES

[1] A. D. Warbhe, R. V. Dharaskar "Blind Method for Image Forgery Detection: A tool for Digital Image Forensics", IJCA, Number 11 (ISBN: 978-93-80866-82-5), 2012.

[2] Hartung F, Kutter M. Multimedia watermarking techniques. Proc. IEEE, July 1999, 87(7): 1079–1107

[3] LUO Weiqi, QU Zhenhua, PAN Feng, HUANG Jiwu "A survey of passive technology for digital image forensics" Front. Comput. Sci. China (2007) , 2(1): 1−11

[4] H. Farid "Image Forgery Detection - A Survey, IEEE signal processing magazine March 2009, pp. 16-25

[5] J. Fridrich, D. Soukal, and J. Lukás, "Detection of copy move forgery in digital images," in Proc. Digital Forensic Research Workshop, Aug. 2003.

[6] A.C. Popescu and H. Farid, "Exposing digital forgeries by detecting duplicated image regions," Dept. Comput. Sci., Dartmouth College, Tech. Rep. TR2004-515, 2004.

[7] G. Li et.al., "A sorted neighborhood approach for detecting duplicated regions in image forgeries based on DWT and SVD," in IEEE Int. Conf. Multimedia and Expo, Beijing, China, 2007, pp. 1750–1753.

[8] W. Luo, et. al., "Robust detection of region-duplication forgery in digital images," in Proc. Int. Conf. on Pattern Recognition, Washington, D.C., 2006, pp. 746–749.

[9] B. Mahdian, S. Saic, "Detection of copy move forgery using a method based on blur movement invariants," Forensic Sci. Int., vol. 171, pp. 180–189, 2007.

[10] A. C. Popescu and H. Farid, "Exposing digital forgeries by detecting traces of re-sampling," IEEE Trans. Signal Processing, vol. 53, no. 2, pp. 758–767, 2005.

[11] A.C. Gallagher, "Detection of linear and cubic interpolation in jpeg compressed images," in Proc. 2nd Canadian Conf. Computer and Robot Vision., Victoria, British Columbia, Canada, vol. 171, 2005, pp. 65–72.

[12] M. Kirchner, "Fast and reliable resampling detection by spectral analysis of fixed linear predictor residue," ACM Multimedia and Security Workshop, 2008, pp. 11–20.

[13] B. Mahdian, S. Saic, "Blind authentication using periodic properties of interpolation," IEEE Trans. Inform. Forensics Security, vol. 3, no. 3, pp. 529–538, 2008.

[14] S. Prasad and K. R. Ramakrishnan, "On resampling detection and its application to image tampering," in Proc. IEEE Int. Conf. Multimedia and Exposition, Toronto, Canada, 2006, pp. 1325–1328.

[15] H. Farid, "Detecting digital forgeries using bispectral analysis," AI Lab, Massachusetts Institute of Technology, Tech. Rep. AIM-1657, 1999.

[16] T.-T. Ng and S.-F. Chang, "A model for image splicing," in Proc. IEEE Int. Conf. Image Processing, Singapore, 2004, vol. 2, pp. 1169–1172.

[17] H. Farid and S. Lyu, "Higher-order wavelet statistics and their application to digital forensics," in Proc. IEEE Workshop on Statistical Analysis in Computer Vision (in conjunction with CVPR), Madison, WI, 2003.

[18] S. Bayram, I. Avcibas, B. Sankur, and N. Memon, "Image manipulation detection with binary similarity measures," in Proc. European Signal Processing Conf., Turkey, 2005.

[19] S. Bayram, I. Avcibas, B. Sankur, and N. Memon, "Image manipulation detection," J. Electron. Imaging, vol. 15, no. 4, p. 41102, 2006.

[20] S. Lyu and H. Farid, "How realistic is photorealistic?" IEEE Trans. Signal Processing, vol. 53, no. 2, pp. 845–850, 2005.

[21] S. Lyu and H. Farid, "Steganalysis using higher-order image statistics," IEEE Trans. Inform. Forensics Security, vol. 1, no. 1, pp. 111–119, 2006.

[22] H. Farid, "Digital image ballistics from JPEG quantization," Dept. Comput. Sci., Dartmouth College, Tech. Rep. TR2006-583, 2006.

[23] H. Farid, "Digital ballistics from jpeg quantization: A followup study," Dept. Comp. Sci., Dartmouth College, Tech. Rep. TR2008-638, 2008.

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

[25] W. Luo, Z. Qu, J. Huang, and G. Qiu, "A novel method for detecting cropped and recompressed image block," in Proc. IEEE Conf. Acoustics, Speech and Signal Processing, Honolulu, HI, 2007, pp. 217–220.

[26] J. Lukas and J. Fridrich, "Estimation of primary quantization matrix in double compressed JPEG images," in Proc. Digital Forensic Research Workshop, Cleveland, OH, Aug. 2003.

[27] A. C. Popescu and H. Farid, "Statistical tools for digital forensics," in Proc. 6th Int. Workshop on Information Hiding, Toronto, Canada, 2004, pp. 128–147.

[28] J. He, Z. Lin, L. Wang, and X. Tang, "Detecting doctored JPEG images via DCT coefficient analysis," in Proc. European Conf. Computer Vision, Graz, Austria, 2006, pp. 423–435.