Imperial Journal of Interdisciplinary Research (IJIR)
Vol-2, Issue-3 , 2016
ISSN : 2454-1362 , http://www.onlinejournal.in

Review on Digital Image Forgeries

Rohini Sharma
Department of Computer Science & Engineering
CGC College of Engineering, Landran, Mohali

Abstract : *With the advent of high-quality digital video cameras and sophisticated video editing software, it is becoming increasingly easier to tamper with digital video. The digital nature of the media files, they can now be easily manipulated, synthesized and tampered in numerous ways without leaving visible clues. As a result, the integrity of image or video content can no longer be taken for granted and a number of forensic-related issues arise. In this paper various detection algorithms and methods has been discussed which detects the authenticated images being tampered.*

1. Introduction

Cameras are regarded as trustworthy devices and photos traditionally imply truth. Nowadays, digital photos have been widely used as historical records and as evidences of real happenings in applications from journalist reporting, police investigation, law enforcement, insurance, medical and dental examination, military and museum to consumer photography. While digital photos are conveniently used, their credibility has been severely challenged due to numerous fraudulent cases involving image forgeries, e.g. the fake results on human stem-cell research. The rapid growth of image editing software's has given rise to large amount of doctored images circulating in our daily lives, generating a great demand for automatic forgery detection algorithms in order to determine the authenticity of a candidate image in a timely fashion.

The fast development of commercial image editing software's such as Adobe Photoshop dramatically increases the amount of doctored photographs. This phenomenon leads to serious consequences, reducing trustworthiness and creating false beliefs in much real-world application. Many image forensic techniques have been proposed during the last decade with the objective to faithfully detect image forgeries. Compared to the authentication based on digital watermarking, forensic techniques can assess the authenticity of an image in a passive and blindway, without resorting to previously embedded information (i.e. the watermark). These techniques make assumption that manipulating an image will probably disturb the intrinsic property, geometrical, physical or statistical, of the authentic image. Therefore, inconsistencies in these properties over the image can be considered as an evidence of tampering.

Digital watermarks have been proposed as a tool to provide authenticity to images, it is a fact that the overwhelming majority of images that are captured today do not contain a digital watermark. And this situation is likely to continue for the foreseeable future. Hence in the absence of widespread adoption of digital watermarks, it is imperative to develop techniques that can help us make statements about the origin, veracity and nature of digital images.
Most work in image forensics in the past two decades has focused on watermarking. In the watermarking paradigm, a unique hidden digital signature needs to be embedded into an image before the image is released. In most cases such insertion must fall below human perception levels so that human eyes cannot detect the inserted signatures. At the receiving end, if the copyright is ever in question, the watermark is extracted and verified to determine the ownership and the authenticity of an image. This active approach, although proven effective in terms of robustness and accuracy, has its fundamental limitations. With the ease of access to image editing tools nowadays, almost everyone can generate tampered images and it is difficult to ensure every image goes through the standard watermarking process. Even if no watermark is extracted from an image, one still cannot claim this image being tampered. Therefore watermarking has limited use in practice. The alternative is to resort to passive approaches. Namely, without assuming any embedded signatures in the image, one looks at the traces inevitably left by the generation or manipulation processes.

1.1 Source identification and Forgery detection

Forgery detection attempts to discover evidence of tampering by assessing the authenticity of the digital media (audio clips, video clips, images etc).

1.2 Tampered data:

A) Publishing business - Newspapers and magazines

Computerized solutions make it possible to verify the authenticity of photographs prior to publishing. The need of automation and high throughput is obvious given the timely nature of news articles and the amount of photographs that have to be processed every day.

B) Criminal justice

Photographs are often presented as court evidence. For this, authenticity verification of every single piece of evidence needs to be solid. Besides expediting the verification process, computerized algorithms can also avoid potential malicious human intervention, thus creating a more objective investigation process.

C) Finance industry

The finance industry can benefit from forensics techniques too, as they have to process and analyze numerous transaction
documents everyday. Financial fraud involves huge monetary loss or gain, therefore there is very little tolerance for miss detection. Such institutions require forensics tools that are both fast and reliable.

1.3 Digital Forensic Analysis

In general, the goal of digital forensic analysis is to identify digital evidence for an investigation. An investigation typically uses both physical and digital evidence with the scientific method to draw conclusions. Examples of investigations that use digital forensics include computer intrusion, unauthorized use of corporate computers, child pornography, and any physical crime whose suspect had a computer. At the most basic level, digital forensics has three major phases:

A) Acquisition

The Acquisition Phase saves the state of a digital system so that it can be later analyzed.
This is analogous to taking photographs, fingerprints, blood samples, or tire patterns from a crime scene. As in the physical world, it is unknown which data will be used as digital evidence so the goal of this phase is to save all digital values. At a minimum, the allocated and unallocated areas of a hard disk are copied, which is commonly called an image.

Tools are used in the acquisition phase to copy data from the suspect storage device to a trusted device. These tools must modify the suspect device as little as possible and copy all data.

B) Analysis

The Analysis Phase takes the acquired data and examines it to identify pieces of evidence.

There are three major categories of evidence:

- **Inculpatory Evidence:** Which supports a given theory
- **Exculpatory Evidence:** Which contradicts a given theory
- **Evidence of tampering**: Which cannot be related to any theory, but shows that the system was tampered with to avoid identification

This phase includes examining file and directory contents and recovering deleted content. The scientific method is used in this phase to draw conclusions based on the evidence that was found.

Tools in this phase will analyze a file system to list directory contents and names of deleted files, perform deleted file recovery, and present data in a format that is most useful.

C) Presentation

The Presentation Phase though is based entirely on policy and law, which are different for each setting. This phase presents the conclusions and corresponding evidence from the investigation. In a corporate investigation, the audience typically includes the general counsel, human resources, and executives. Privacy laws and corporate policies dictate what is presented. In a legal setting, the audience is typically a judge and jury, but lawyers must first evaluate the evidence before it is entered. In order to be admissible in a United States legal proceeding, scientific evidence must pass the so-called “Daubert Test”, which stems from the U.S. Supreme Court’s ruling in Daubert vs. Merrell Dow Pharmaceuticals (1993)

1.4 Digital media to be altered and manipulated as:

A) Cloning: One of the most common image manipulations 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.

B) Resampling: To create a convincing composite, it is often necessary to resize, rotate, or stretch portions of an image. 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.

C) Splicing: A common form of photographic manipulation is the digital splicing of two or more images into a single composite. When performed carefully, the border between the spliced regions can be visually imperceptible.

2. Various Models to be used:

1 Image forgery using codebook :
Codebook method is used to detect the image forgery. The codebook is generated from the set of training images, by extracting the SIFT features and clustering. The centroids are taken to generate the codebook to improve the accuracy. The work consists of five steps- SIFT feature Extraction, Clustering, Codebook Generation, Tampering detection and Locating the image forger

A) SIFT Feature Extraction :

In feature extraction only stable points are extracted. Edge points and the points that are sensitive to noise are eliminated. The SIFT extraction method is categorized into four steps- Scale space extrema detection, Key point localization, Orientation assignment and Key point descriptor.

Scale space extrema searches over all scales and image locations and constructs scale space. To identify potential interest points that are invariant to scale and orientation, a difference of Gaussian function is computed. There are large no of key points in key point localization. In order to remove the unwanted keypoints orientation assignment is done. After selecting the orientation, the feature descriptor is computed.

B) Clustering:
In order to partition n observations into k clusters, k mean clustering is used and each observation is belongs to the cluster with the nearest mean.

C) Code book generation:
The set of training images are generated by codebook and the features were extracted using SIFT algorithm. After the extraction of the features clustering is done and the centroids are extracted from the clusters to generate the codebook. For Test image code book is also generated using sift algorithm.

D) Tampering detection:

The test image codebook is compared with the training image's codebook and geometric manipulation will be detected.

1.1 Locating the image forgery :

The image forgery has been located by extracting the gradient features and gradient of a function F(x,y) is calculated as:

$$F=\frac{\partial F}{\partial x}i+\frac{\partial F}{\partial y}j$$

2 Image forgeries by measuring inconsistencies of blocking artifact :

When a digital forgery is created the resultant tampered image may inherit different kind of artifacts from different sources and if these inconsistencies detected, could be used to detect image integrity [2]. Image forgery creation process also changes the blocking artifact because the blocking artifacts of the affected blocks will change a lot by tampering operations such as image splicing, resampling.

Blocking artifact is estimated as:

$$B(i)=\sum_{k=1}^{64}|D(k)-Q(k)\,\mathrm{round}(\frac{D(k)}{Q(k)})|$$

where B(i) is the estimated blocking artifact for the testing block i, and D(k) is the DCT coefficient at position k. Q(1:64) is the estimated DCT quantization table.

3 Forgery detection Using Robust Matching:

The work has been focused on the detection of a particular type of forgery, the copy-move attack, where a part of an image is cloned elsewhere in the same image, usually to conceal an important feature [4]

For uncompressed images, matching is carried out between blocks of size BxB to detect for exact replicas. To extend this idea to images saved in lossy JPEG format, instead of directly matching the pixel representation of each BxB block, the authors use a robust representation consisting of quantized DCT coefficients [3]

4 Forgery detection using Harris interest points and SIFT descriptors:

Copy- moved forgery is one type of forgery in which one part of the image is itself copied and pasted into another part of the same image. This method is mainly focused on detection of copy-moved forgery based on harris corner detector and sift descriptors.

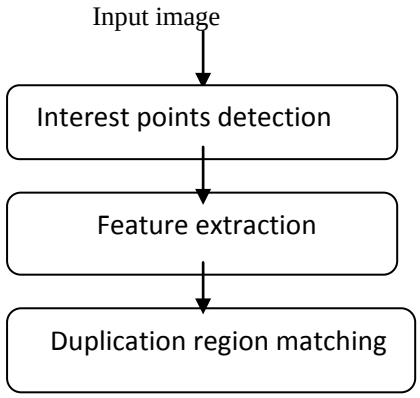

Fig1: A schematization of the whole system

4.1 Interest Point Detection: The fast, robust and rotation invariant Harris detector is used which uses the autocorrelation function to determine locations where the change of signal in one or two directions. A matrix is related to auto-correlation function is computed as:

$$C(x, \sigma_I, \sigma_D) = \sigma_D^2 G(x, \sigma_I) * \begin{pmatrix} I_x^2(x,\sigma_D) & I_x I_y(x,\sigma_D) \\ I_x I_y(x,\sigma_D) & I_y^2(x,\sigma_D) \end{pmatrix}$$

where σ_D is the derivation scale, σ_I is the integration scale , G is the Gaussian and L is the image smoothed by a Gaussian kernel

This matrix has two Eigen values that are the principal curvatures of the auto-correlation function. When the two eigenvectors are very small then there is no structure exists. If one is large and another one is small, there is an edge like structure. If both of them are very large and distinct, there is a corner like structure [6]. Edges and interest points can be computed based on:

$$\det(C) - a.\ \text{trace}^2(C) < T_E$$

$$\det(C) - a.\ \text{trace}^2(C) < T_C$$

Edges are computed based on these equations, where α is the coefficient of the Harris function and TE is the threshold of the Harris function (TE < 0). The edge detection is carried out at the first scale. Interest points can be detected by using eq. (3), TC is the threshold for interest points (TC > 0).

4.2 Feature Extraction: For feature extraction SIFT descriptors are used. The feature descriptor is computed as a set of orientation histograms on 4 x4 pixel neighbourhoods.

Conclusion: Since image forensics is a real world problem, a good tampering detection system should meet realistic requirements. In this paper various tampering detection tools can be applied to a wide variety of images. The proposed work is focused on robustness against possible attacks and to design a system with reliable features, extraction processes and authenticity checking components so that tampered images can still be successfully detected even when extra attacks have been applied.

References

1. E. Agnes *et al*, "A Forensic Method for Detecting Image Forgery Using Codebook", *International Journal of Advanced Research in Computer Science and Software Engineering*, Vol. 3, No. 3, 2013

2. Shuiming Ye *et al*, "Detecting digital image forgeries by measuring inconsistencies of blocking artifact", *IEEE*, 2007

3. Tran Van Lanh *et al*, "A survey on digital camera image forensic methods", *IEEE*, 2007

4. J. Fridrich *et al*, "Detection of Copy-Move Forgery in Digital Images", *Proc. of DFRWS* , 2003

5. H. Farid, "Digital Image Ballistics from JPEG Quantization",Technical Report, Dartmouth College, Computer Science, 2006

6. B.L. Shivakumar, "Automated Forensic Method for Copy-Move Forgery detection based on Harris Interest Points and SIFT Descriptors" *International Journal of Computer Applications*, vol. 27, No. 3, 2011

7. Weiqi Luo *et al*, "A novel method for detecting cropped and recompressed image block" , *IEEE*, 2007

8. Vincent Christlein, "An Evaluation of Popular Copy-Move Forgery Detection Approaches", *IEEE transactions on information forensics and security*, 2012

9. Hany Farid, "Exposing Digital Forgeries From JPEG Ghosts", *IEEE transactions on information forensics and security*, Vol. 4, No.1 , 2009

10. Weiqi Luo, "JPEG Error Analysis and Its Applications to Digital Image Forensics" *IEEE transactions on information forensics and security*, Vol. 5, No. 3, 2010

11. Xunyu Pan, "Region Duplication Detection Using Image Feature Matching", *IEEE transactions on information forensics and security*, vol.5, no. 4 , 2010

12. Matthew C. Stamm, "Forensic Detection of Image Manipulation Using Statistical Intrinsic Fingerprints", IEEE *transactions on information forensics and security*, Vol. 5, No. 3, 2010