Passive Approaches for Detecting Image Tampering: A Review

Fatma Salman Hashem*, Ghazali Sulong

UTM-IRDA Digital Media Centre (MaGIC-X), Faculty of Computing, Universiti Teknologi Malaysia, 81310 UTM Johor Bahru, Johor, Malaysia

Corresponding author: eng_fatmas@yahoo.com

Abstract

This paper defines the presently used methods and approaches in the domain of digital image forgery detection. A survey of a recent study is explored including an examination of the current techniques and passive approaches in detecting image tampering. This area of research is relatively new and only a few sources exist that directly relate to the detection of image forgeries. Passive, or blind, approaches for detecting image tampering are regarded as a new direction of research. In recent years, there has been significant work performed in this highly active area of research. Passive approaches do not depend on hidden data to detect image forgeries, but only utilize the statistics and/or content of the image in question to verify its genuineness. The specific types of forgery detection techniques are discussed below.

Keywords: Detecting; image; forgery; passive; tampering; techniques; tools

© 2015 Penerbit UTM Press. All rights reserved.

1.0 INTRODUCTION

Maliciously manipulate, and tamper digital images without leaving any obvious clues became very easy with the widely available, easy to use and extremely powerful digital image processing tools such as Photoshop. As a result, there is a rapid increase of the digitally manipulated splicing in mainstream media and on the Internet. This trend indicates serious vulnerabilities and decreases the credibility of the digital images. Therefore, developing techniques to verify the integrity and the authenticity of the digital images became very important, especially considering the images presented as evidence in a court of law, as news items, as a part of a medical record, or as a financial document. In this sense, image tamper detection is one of the primary goals in image forensics.

Fake images have become widespread in society today. The accessibility to powerful simple to use image editing computer software to end users helps make the job of manipulating image incredibly easy. One can find forged images used to sensationalize news, spread political propaganda and rumors, introduce psychological bias, etc. in all forms of media. Claims of image tampering are common in scandals and controversies. As the credibility of images suffers, it is necessary to devise techniques in order to verify their authenticity and trustworthiness of digital images [1].

Before discussing the various types of forgery detection techniques in existence, it is important to know what forgeries they are dealing with. To this end, we classify forgeries into five major categories. Figure 1 depicts the classification of image forgery approaches as follows:

1.1 Splicing

One of the types of image tampering is splicing or composition or photomontage. In such a forgery, elements from multiple images are often juxtaposed in a single image to convey an idea that could not have been conveyed by any of the original images. Such an idea usually does not reflect reality, and so such spliced images can be very damaging. Examples of some prominent image forgeries are shown in Figure 2.
Figure 2 shows a hoax image that surfaced right after flooding in Puerto Rico caused due to Hurricane Irene in 2011. The inset image shows a shark swimming down a flooded street, but is a hoax with the shark likely digitally inserted from the larger image published in a 2005 issue of Africa Geographic.

Such splicing can usually be detected by searching for the splicing boundary (or the effect of the splicing on image statistics), or by considering the directions of the light incident on surfaces in the image. Other abnormalities such as inconsistent demosaicking or chromatic aberration may also be used to determine the inauthenticity of such images. Inconsistencies in lighting or blurred splicing boundaries can be used to expose the above images as fake, if the light direction can be correctly estimated or if the splicing boundary can be correctly detected respectively.

1.2 Copy-Move

Another common type of image forgery is the copy-move (or region duplication or cloning) forgery. In this type of forgery, regions from the same image are copied and pasted (with possible transformations) in the same image. This is usually done with the intent of hiding certain content present in the original image or duplicating certain content not actually present in the image. Some examples of copy-move forgeries are shown in Figure 3.

The left panel in Figure 3 shows an image that was used in a political advertisement for the presidential campaign of George W. Bush in 2004. It was created by copying and pasting a set of soldiers from the center of the image, over Mr. Bush, thus creating a more patriotic image of a child waving a flag in front of a sea of soldiers.

Copy-move forgeries are usually detected by searching for matching regions in the image, although recent research has taken a more SIFT-based approach, concentrating on matching key points (as in object detection) rather than blocks, in order to allow for various image transformations that can be used to create more convincing forgeries.

1.3 Image Retouching

Image Retouching can be regarded to be the much less harmful type associated with digital image forgery. Image retouching does not really considerably modify an image, however rather, improves or even decreases certain feature of an image (see Figure 4). This method is well-liked by magazine photo editors. It can probably be asserted just about all magazine cover would utilize this technique to improve particular features of the image so that it can be more attractive; disregarding the truth that this kind of enhancement is morally wrong.

1.4 Image Processing Operations

The types of forgeries described in this subsection and the next do not necessarily fall into the traditional definition of image tampering i.e. there is no addition or hiding of information. However, as in the case of the image tampering discussed in this subsection, the reality portrayed by an image processing operationcan be distorted at a psychological level. Samples of such tampering are shown in Figure 5.

Figure 5 shows an example of an image from the Charlotte Observer of July 2006 where the hue of the sky was changed from a dull brownish-gray to a fiery orange to yield a more striking effect. It is debatable if this constitutes actual tampering because the content of the image remained unaffected. However, the Observer considered the altering of image colors as a violation of its photo policy and fired the staff photographer who implemented the change, thus taking the stand that this was indeed a case of image tampering.
It is possible to detect various types of filtering and tonal operations in images by studying the characteristics of images subjected to these operations, and then looking for similar characteristics in the image under examination.

1.5 False Captioning

The last type of image forgery that we discuss is probably the least like the other three types seen so far. In false captioning it is possible that the content of the image is not touched at all, but the caption of the image, which provides context, is changed from the actual situation with intent to mislead the viewer or reader. Examples of images where false captioning has been used are shown in Figure 6. Apart from actual image captions, other informative elements like metadata (which can provide geographical or temporal context), may be tampered without changing the image content. Such tampering also falls within the category of false captioning.

Figure 6 shows US President George W. Bush carrying a turkey on a surprise visit to US troops in Iraq, during Thanksgiving in 2003. There was a spike in the president’s approval ratings after this trip, which had the above image as its most widely circulated one. However, it was later revealed that the particular turkey in the image was not meant for US troops at all, and the President has just picked it up. Thus, implied false context again distorted reality. Although not much work has been done in this regard, rudimentary AI-based techniques have been proposed to detect wildly incorrect or implausible captioning in images. Certain techniques also allow for identification of source cameras by using camera-specific characteristic like color filter arrays, sensor noise patterns, etc. The camera make and model is often recorded in the image metadata and so, these techniques can be applied to verify these metadata elements. However, color filter array patterns often tend to overlap between different models of cameras, and identifying a camera from its sensor pattern noise requires access to the specific camera.

2.0 PASSIVE APPROACHES

Passive, or blind, approaches for detecting image tampering are regarded as a new direction of research. In recent years, there has been significant work performed in this highly active area of research. Passive approaches do not depend on hidden data to detect image forgeries, but only utilize the statistics and/or content of the image in question to verify its genuineness. General surveys of this field may be found in [2, 3, 4]. The specific types of forgery detection techniques are discussed below.

2.1 False Captioning

The most obvious way of identifying image forgeries has to be to look at the pixels constituting the image. Various techniques either directly or indirectly utilize possible correlations that occur between pixels as a result of tampering.

One of the most common types of image forgeries is known as the copy-move (or copy-paste or cloning) forgery. In this particular type of forgery, one region of the image is concealed by using another region from the same image to cover it. The copied region may be subjected to some image processing operations in order to make cloning difficult to detect visually.

As the cloned regions can be of any shape and location, it is infeasible to search all possible image locations and sizes. Dividing the image into blocks and applying a brute force search is also computationally prohibitive, besides not being robust to simple anti-forensic measures like noise addition. Therefore, initial work in this direction focused on better representing the image blocks. Two of the most prominent techniques that emerged [5] and [6] use the Discrete Cosine Transform (DCT) and Principal Components Analysis (PCA) respectively, in order to efficiently search for matching blocks in the image. Besides, they are also robust to minor variations in the image due to additive noise or lossy compression. Recently, other techniques [7, 8] have emerged which are robust to additional transforms like scaling, rotation and contrast changes. Although these methods can detect cloning in an image, they are still quite computationally intensive, and tend to produce a lot of false positives. Moreover, a human interpretation of the output is necessary for these techniques.

In order to create convincing forgeries, it is often necessary to apply various image processing operations to the image. The detection of such operations indicates that the image is not in the same form as captured, and may have been tampered with. For instance [9, 10], propose methods to detect median filtering in images. [9] propose a method to detect if an image has been subjected to gamma correction, while [10] check for contrast enhancement, including histogram equalization. Another characteristic which can be exploited to detect composite images is the presence of resampling. In order to create convincing composites, it is often necessary to resize certain objects. Such resizing introduces certain unnatural correlations between neighboring pixels, which, if detected, can indicate possible tampering in the image. Examples of techniques that utilize traces of resampling to detect forgeries are [10, 11, 12, 13]. Although the work of [11] improves upon previous related work by considering phase-dependent resampling prediction rather than linear resampling prediction, strong JPEG compression tends to create artifacts which can hide traces of filtering and resampling, which is a limitation of these techniques.

Even in the absence of resampling, creating composites (also known as splicing) gives rise to certain abnormalities at the splicing boundary. If the splicing is done carefully, then the boundary between the regions can be visually imperceptible. However, higher-order Fourier statistics are disrupted by splicing, which can hint at the presence of splicing [14]. Techniques based on this principle utilize the bispectrum to analyze higher-order correlations between frequencies. As with the above techniques, compression artifacts tend to limit the applicability of these methods.

It is important to note that perceptually meaningful images are not collections of pixels having random intensities. For example, pixels tend to display correlations with their neighbors, corresponding to objects in the images. There are other statistical properties which can be examined to verify image authenticity. Methods that examine such statistical properties make use of statistical moments from a wavelet decomposition of the [15], and local co-occurrence characteristics in image bit-planes [16].
Such techniques have found wide application, for instance, in distinguishing between photorealistic and photographic images [17, 18] (with the latter improving performance by using an increased number of features along with boosted feature selection to manage computational complexity). These techniques typically require uncompressed images or JPEG images with a high quality factor to be successful.

2.2 Format-based Techniques

One of the most common image formats used today is the JPEG lossy compression format. This is based on representing the image as Discrete Cosine Transform (Discrete Cosine Transform (DCT)) blocks, and quantizing the resulting coefficients. This quantization is the source of the lossy compression in this technique. Such compression gives rise to certain artifacts which can be exploited to detect tampering.

The manner in which the DCT coefficients in each block are quantized is determined by a quantization table. The quality of the image and its size is determined by the quantization table, and these tables tend to differ between camera manufacturers. This difference between tables may be exploited to perform a forensic analysis on the image to determine its source camera [19, 20]. The quantization table may be available in the header of the JPEG image, or may be determined blindly as in [21]. One of the limitations of this technique is that the quantization table used within a camera often depends on the quality setting at which the image is captured. Moreover, with the incredibly large number of digital cameras available commercially, there is bound to be some overlap of the quantization tables. Hence, such a method alone cannot provide conclusive evidence of the source of an image.

Another case to consider is that of a JPEG image being tampered with and resaved. This results in the JPEG image being compressed twice. Such double compression gives rise to certain artifacts not visible in singly compressed images. Techniques presented in [22, 23] (with addressing the difficult case of the quantization matrices for both compressions being the same) describe methods to detect these artifacts, and provide evidence of manipulation. However, it is to be noted that detection of such double compression does not necessarily imply malicious intent. For instance, it is entirely possible to resave an image with a lower quality factor for faster transmission over a communication network.

As each block is transformed and quantized independently from other blocks in a JPEG image, horizontal and vertical edges often appear between blocks as artifacts. When a JPEG image is tampered with, certain aberrations may result in these artifacts. Detection of these aberrations as described in [24, 25, 26] can detect manipulations and manipulated regions in images.

Obviously, these techniques do not work for non-JPEG images, as they rely on the artifacts introduced by the JPEG process. Also, knowing that an image has been resaved is often not enough, when it is required to know the specific tampering that has occurred.

2.3 Camera-based Techniques

It is often necessary to establish the source of an image - in a court of law, for example even if the image has not been specifically tampered with. As cameras are not perfect imaging systems, there are certain artifacts present in cameras which can be used to associate an image with a specific camera. Various techniques exist to facilitate this form of image forensics.

Camera lens aberrations can often be used to identify the source camera, and even to detect image tampering. For instance, [27] proposes a technique to identify the lens with which an image was captured by analyzing the artifacts in the image resulting from dust specks on the lens. It is important to note that a negative result from this technique does not necessarily mean that an image could not have originated from a camera with a particular lens, because the lens may have been cleaned of the dust. Another lens aberration that can be employed is lateral chromatic aberration [28, 29, 30]. This results from the tendency of light of different wavelengths to be bent to different extents by the lens resulting in an expansion or contraction of the color channels with respect to each other. If another object is added into the image, it is likely that the expansion/contraction pattern will be disturbed allowing for the tampered region to be detected. Such a technique works well only for non-compressed or non-uniform parts of the image, although [31] reports improved performance by using a more visible, but also more region-dependent, purple fringing aberration as a generalization of local chromatic aberration.

In order to keep costs low, most digital cameras often have a single CCD or CMOS sensor in order to capture color images. This is accomplished by using a color filter array (CFA). Each sensor element records only one of red, green or blue color channel samples and interpolates the missing two from the neighboring samples. This process is known as demosaicking, and the particular algorithm adopted (bilinear, bicubic, adaptive, etc.) can be used to distinguish between cameras. The particular type of interpolation can be identified from the statistical periodic correlations introduced between subsets of pixels in each color channel [12, 32, 33]. The technique of [33] generalizes the approaches of previous techniques and performs well in correctly identifying a large number of demosaicking algorithms. Deviations from the periodic correlation pattern can be used as evidence of global or local tampering.

As camera sensors tend to be linear, a linear relationship is expected between the intensity of light incident on each sensor and the resulting intensity value of each pixel. However, in order to enhance the final image, cameras often apply a pointwise nonlinearity. This nonlinear mapping can be estimated and discrepancies in the mapping can be used to detect tampering, as described in [31, 34]. As with many other techniques, compression artifacts can make it very difficult to localize forgeries.

Images captured with digital cameras are subjected to a number of image processing operations between the camera sensor and memory, such as white balancing, contrast enhancement, filtering, etc. It is possible to model these operations and detect if an image has undergone any subsequent processing [35]. Camera sensors also contain various sources of noise, such as dark current noise and photosensponsenonuniformity (PRNU) noise. The latter has been shown to be distinct for specific sensors [36], and can be used to identify the specific camera with which the image was captured. However, this technique requires access to the camera or a large number of images taken with the camera.

2.4 Physics-based Techniques

One of the biggest challenges in creating a convincing spliced image is to match the light-source directions of the images being combined. Differences in lighting can be used as evidence of tampering in an image. The techniques in this subsection necessitate human interpretation of the output because of their nature.

The lighting direction can be estimated at various points in an image from the two-dimensional surface normals at the occluding object boundary [37]. By assuming Lambertian surfaces with constant reflectance values and a point light source at infinite distance, a set of equations can be solved for the lighting directions and ambient light terms. Inconsistencies in lighting can, and have, been used to expose various forgeries.
As 3-D surface normal are difficult to estimate from an image in general, there still remains an ambiguity in the light source direction in the above technique. [37] helped remove this ambiguity in certain cases where the image contains people, and their eyes are clearly visible. This is accomplished by using the specularities resulting in the eye from the light source. The 3-D normals are determined from the 3-D model of the human eye. This estimated direction can be compared for various people in the image, or with the estimated direction as explained above.

In practical scenarios, multiple light sources may be present, instead of a single dominant one as assumed so far. [36] Discuss how a lowparameter representation of such a complex lighting environment may be achieved. By assuming the light striking a Lambertian surface to be a weighted sum of spherical harmonic functions, the light source directions in two dimensions may be estimated and checked for consistency across the image.

In specific cases, computer graphics techniques have been used to simulate the physical conditions of the scene depicted and check its feasibility [39, 40]. However, considerable human intervention is needed in the creation of the scene models, and it is difficult to generalize such techniques. An attempt at generalization is found in [41] where morphable 3-D models of human faces are used to check for lighting direction consistency. Although such a technique reduces the amount of human input needed, it is still limited to a very specific subset of image forgeries.

2.5 Other Techniques

The above subsections outline the major classes into which forgery detection techniques tend to fall. There are other techniques as well, which do not neatly fall into the above categories. These usually rely on semantic or geometric knowledge of the scene depicted in the image.

Sometimes, it may so happen that the content of a digital image is another digital image, either being viewed on a screen or having been printed as a hard copy. This is known as image recapturing and although not technically a forgery in its own right, it can prove to be an obstacle for camera and image forensics techniques. Various approaches [42, 43] have been formulated to deal with the detection of such recapturing. However, they depend on the surface characteristics of the source of the recaptured image (LCD screens [42], printing paper [42, 43], etc.) which are difficult to identify with common image post processing operations such as JPEG compression and contrast enhancement.

In an un tampered image, the projection of the camera center onto the image plane the principal point - is located near the center of the image. In a spliced image, the location of objects can differ from their locations in the individual original images. In copy-move forgeries, objects or regions are often translated across the image. Both such operations result in the principal point moving proportionally. [43] Describe a technique to estimate the principal point of a camera from certain planar geometric shapes present in the image, and utilize discrepancies in the estimate to provide evidence of tampering.

It is often difficult to discern certain details in an image because of the angle at which the shot was captured. In [44], various projective geometric tools are reviewed. These tools allow for metric measurements to be made from a single image, under certain assumptions. These depend on knowledge of polygons, vanishing points and coplanarity of circles in order to remove planar distortions that enable metric measurements to be made on the plane.

[38] Propose a method that utilizes the differences in geometry between authentic and computer-generated images in order to distinguish between the two. These differences include the non-uniform nature of real objects versus the relatively smooth surfaces of computer-generated objects, and lighting assumptions such as isotropy that are used in creating photorealistic images, but may not be valid for photographic images. The technique for identifying these differences is inspired by the physics of the image formation process, but requires high-quality images.

An important issue in image forgery detection is the semantic content of the image. In [41], perceptually meaningful regions of an image are found and, using AI techniques, checked against characteristics of such regions to detect false captioning. However, the work in this direction is very preliminary and can only detect extremely incorrect captions.

3.0 SUMMARY

This paper discussed the current state of research in terms of digital image forensics. While digital watermarking has been the method of choice to safe-guard one’s images from manipulation and to secure a copyrightable image, it has been difficult to determine if an image of unknown origin is authentic. Several techniques exist that touch the surface of the subject. These hold some sound results, as previously discussed, but further analysis is needed to determine the best and most efficient method to detect image forgeries. A survey of a recent study is explored including an examination of the passive approaches in detecting image tampering. This area of research is relatively new and only a few sources exist that directly relate to the detection of image forgeries. Passive, or blind, approaches for detecting image tampering are regarded as a new direction of research. In recent years, there has been significant work performed in this highly active area of research. Passive approaches do not depend on hidden data to detect image forgeries, but only utilize the statistics and/or content of the image in question to verify its genuineness.

References


