ISSN No. 0976-5697

Volume 5, No. 8, Nov-Dec 2014

International Journal of Advanced Research in Computer Science

RESEARCH PAPER

Available Online at www.ijarcs.info

An Approach towards Image Forgery Detection

Mr. S. B. Lanjewar
Asst. Prof.
Dept. of Computer Science & Engineering
Dr. Babasaheb Ambedkar College of Engg. & Research
Nagpur, India

Mr. P. A. Khaire
Asst. Prof.
Dept. of Computer Science & Engineering
S. B. Jain College of Engineering Research & Management
Nagpur, India

Mr. R. Meshram
Asst. Prof.
Dept. of Computer Technology
Government polytechnic
Nagpur, India

Abstract: As one of the most successful applications of image analysis and understanding, digital image forgery detection has recently received significant attention, especially during the past few years. At least two trend account for this: the first accepting digital image as official document has become a common practice, and the second the availability of low cost technology in which the image could be easily manipulated. Even though there are many systems to detect the digital image forgery, their success is limited by the conditions imposed by many applications. For example, detecting duplicated region that have been rotated in different angles remains largely unsolved problem. In an attempt to assist these efforts, this project surveys the recent development in the field of Copy- Move digital image forgery detection.

Keywords: passive techniques, active techniques, cloning, splicing, re-sampling

I. INTRODAUCTION

The field of digital forgery had drawn the attention researcher worldwide. The problems being highlighted in this domain are digital forgeries of social impacts, detection techniques, and prevention techniques. The digital forgeries have many perspectives and implications on social, legal, technical, intelligence, investigative mechanisms, security, managerial issues. Prof. Hanyfarid has lately drawn attention of researcher worldwide by reporting digital image forgery problem [1]. Due to rapid advances and availabilities of powerful image processing software's, it is easy to manipulate and modify digital images. So it is very difficult for a viewer to judge the authenticity of a given image. For digital photographs to be used as evidence in law issues, it is necessary to check the authenticity of the image.

Image Forgery Detection is probably one of the most interesting functions under Digital Image Forgery due to its application which is generally much closer to the public compared to the other two functions. It deals with techniques or algorithm to detect traces of digital image tampering. The availability of any of these traces is proof that an image has been tampered with. However, the absence of these traces does not indicate that the image is authentic or has not been modified.

The other side of forgery are those who perpetuate a forgery for gain and prestige they create an image in which to dupe the recipient into believing the image is real and from this be able to gain payment and fame.

Three type of forgery can be identified:

1) An image that is created using graphical software
2) An image where the content has been altered
3) An image where the context has been altered

Using graphical software is one method in which a forged image can be created.

It needs the creator to especially skilful in ensuring that the image they are creating is realistic, for example, that the fall of light on objects in an image is consistent right across the image, that shading is consistent, the absorption of light. An image created using this method can take some time to develop. Creating an image by altering its content is another method. Duping the recipient into believing that the objects in an image are something else from what they really are.

The image itself is not altered, and if examined will be proven as so. This method is where the context of the image is altered. Objects are be removed or added, for example, a person can be added or removed. The easiest way is to cut an object from one image and insert it into another image – image editing software makes this a simple task.

The objective of this project is to develop a new forgery detection method that does not rely upon digital watermarks. It is an area which has been researched by a number of scientists, but most notably - Alin C. Popescu and Hany Farid who developed a number of techniques[1][2][3].

Techniques include:

- How light falls on objects within an image
- Whether there are any areas of an image which has been “cloned” or “copied”
- Looking for evidence of “retouching” in an image by detecting whether the pattern in the pixels have been destroyed
- Core Areas of Research

- Developing a new technique to detect changes and alterations to an image and does not rely on watermarks
- Evaluate the algorithms used for image manipulation and to group them based upon their degree of manipulation. It is also important to consider image compression of image files in this area of work
- An evaluation and comparison of the existing forgery detection techniques, and to carry out the evaluation along with defining a new grouping structure for forgery detection techniques
- To determine whether a given digital image has undergone any form of modification or processing after it was initially captured.
- A common forgery, namely copy-move, in which the copied portion is a textured region from the same image.

Other types of forgeries such as image splicing (putting multiple images together to create a forged image) have not been explored so much.

Digital image forgery is a growing problem in criminal cases and in public course. Currently there are no established methodologies to verify the authenticity and integrity of digital images in an automatic manner. Detecting forgery in digital images is an emerging research field with important implications for ensuring the credibility of digital images.

In order to detect forgery in images, we have applied passive techniques, specifically pixel –based techniques, by which a forged image will be detected.

II. METHODS

Active Techniques

In active approach, the digital image requires some pre-processing such as watermark embedding or signature generation at the time of creating the image. Rely on pre-registration or pre-embedded information and they have not been thoroughly researched.

Digital Watermarking

Digital watermarking is the process of computer-aided information hiding in a carrier signal; the hidden information should, but does not need to contain a relation to the carrier signal. Digital watermarks may be used to verify the authenticity or integrity of the carrier signal or to show the identity of its owners. It is prominently used for tracing copyright infringements and for banknote authentication. Like traditional watermarks, digital watermarks are only perceptible under certain conditions, i.e. after using some algorithm, and imperceptible anytime else [5].

If a digital watermark distorts the carrier signal in a way that it gets perceivable, it is of no use. Traditional Watermarks may be applied to visible media (like images or video), whereas in digital watermarking, the signal may be audio, pictures, video, texts or 3D models. A signal may carry several different watermarks at the same time. Unlike metadata that is added to the carrier signal, a digital watermark does not change the size of the carrier signal. The needed properties of a digital watermark depend on the use case in which it is applied. For marking media files with copyright information, a digital watermark has to be rather robust against modifications that can be applied to the carrier signal[8][9]. Instead, if integrity has to be ensured, a fragile watermark would be applied.

Both steganography and digital watermarking employ steganographic techniques to embed data covertly in noisy signals. But whereas steganography aims for imperceptibility to human senses, digital watermarking tries to control the robustness as top priority. Since a digital copy of data is the same as the original, digital watermarking is a passive protection tool. It just marks data, but does not degrade it nor controls access to the data. One application of digital watermarking is source tracking. A watermark is embedded into a digital signal at each point of distribution. If a copy of the work is found later, then the watermark may be retrieved from the copy and the source of the distribution is known. This technique reportedly has been used to detect the source of illegally copied movies

It is not possible to determine whether the watermark has been inserted after manipulation. Watermarks can be fragile, and when a watermark image has been compressed with using a compressing algorithm like JPEG, they are destroyed. Extract a watermark manipulate the image and then reinsert the watermark, which itself can be modified during insertion therefore making the technique unreliable.

Detecting areas where an image has been manipulated. Determining whether a manipulation is innocent, such as JPEG Compression and sharpening, from those which are malicious, such as adding or removing parts to an image. This is where watermarks have a difficulty – they are unable to determine this difference.

Digital Signature

A digital signature or digital signature scheme is a mathematical scheme for demonstrating the authenticity of a digital message or document. A valid digital signature gives a recipient reason to believe that the message was created by a known sender, and that it was not altered in transit. Digital signatures are commonly used for software distribution, financial transactions, and in other cases where it is important to detect forgery or tampering

Digital signatures employ a type of asymmetric cryptography. For messages sent through a non-secure channel, a properly implemented digital signature gives the receiver reason to believe the message was sent by the claimed sender. Digital signatures are equivalent to traditional handwritten signatures in many respects, but properly implemented digital signatures are more difficult to forge than the handwritten type. Digital signature schemes in the sense used here are cryptographically based, and must be implemented properly to be effective.

Digital signatures can also provide non-repudiation, meaning that the signer cannot successfully claim they did not sign a message, while also claiming their private key remains secret; further, some non-repudiation schemes offer a time stamp for the digital signature, so that even if the private key is exposed, the signature is valid[7][10].

The techniques which are capable of detecting tampering in images from any camera, without relying on watermarks or specialized hardware are passive techniques. Instead of watermarks, these techniques assume that images possess certain regularities that are disturbed by tampering.

Image forensic tools:-

1. Format-Based
2. Camera-Based
3. Physics-Based
4. Geometric-Based
5. Pixel-based

Format-based

The first rule in any forensic analysis must surely be "preserve the evidence." In this regard, 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. I describe three forensic techniques that detect tampering in compressed images, each of which explicitly leverages details of the JPEG lossy compression scheme.

Camera-based

Grooves made in gun barrels impart a spin to the projectile for increased accuracy and range. These grooves introduce somewhat distinct markings to the bullet fired, and can therefore be used to link a bullet with a specific handgun. In the same spirit, several image forensic techniques have been developed that specifically model artifacts introduced by various stages of the imaging process. I describe four techniques for modeling and estimating different camera artifacts. Inconsistencies in these artifacts can then be used as evidence of tampering [8].

Physics-based

Consider the creation of a forgery showing two movie stars, rumored to be romantically involved, walking down a sunset beach. Such an image might be created by splicing together individual images of each movie star. In so doing, it is often difficult to exactly match the lighting effects under which each person was originally photographed. I describe three techniques for estimating different properties of the lighting environment under which a person or object was photographed. Differences in lighting across an image can then be used as evidence of tampering.

Geometric-based

In authentic images, the principal point (the projection of the camera center onto the image plane) is near the center of the image. When a person or object is translated in the image, the principal point is moved proportionally. Differences in the estimated principal point across the image can therefore be used as evidence of tampering. The authors described how to estimate a camera's principal point from the image of a pair of eyes (i.e., two circles) or other planar geometric shapes. They showed how translation in the image plane is equivalent to a shift of the principal point. Inconsistencies in the principal point across an image can then be used as evidence of tampering[2].

Pixel-based

The techniques which are capable of detecting tampering in images from any camera, without relying on watermarks or specialized hardware are passive techniques.

Instead of watermarks, these techniques assume that images possess certain regularities that are disturbed by tampering. The legal system routinely relies on a range of forensic analysis ranging from forensic identification (Deoxyribonucleic acid (DNA) or fingerprint) to forensic deontology (teeth), forensic entomology (insects), and forensic geology (soil). In the traditional forensic sciences, all manner of physical evidence is analyzed. In the digital domain, the emphasis is on the pixel—the underlying building block of a digital image. I 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 [2].

We are going to implement the pixel based techniques for the purpose of detecting the forgery in images. The following are the various pixel based techniques:

Cloning

Perhaps 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. 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 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.

Re-sampling

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 re-sampling 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 [6]. That is, across the entire re-sampled signal, each even sample is precisely the same linear combination of its adjacent two neighbors. In this simple case, a re-sampled signal can be detected by noticing that every other sample is perfectly correlated with its neighbors. His correlation is not limited to up-sampling by a factor of two. A large range of re-samplings introduces similar periodic correlations. If the specific form of the re-sampling 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.

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 [10].

Statistical

There are a total of 256*n*2 possible 8-b gray-scale images of size *n* 3 *n*. With as few as *n*510 pixels, there are a whopping 10240 possible images (more than the estimated number of atoms in the universe) [9]. 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. This decomposition splits the frequency space into multiple scale and orientation sub bands. The statistical model is composed of the first four statistical moments of each wavelet sub band and higher-order statistics that capture the correlations between the various sub bands. Supervised pattern classification is employed to classify images based on these statistical features. 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 is 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.

III. RESULTS AND DISCUSSION

A standard test image is a digital image file used across different institutions to test image processing and image compression algorithms. By using the same standard test images, different labs are able to compare results, both visually and quantitatively. The images are in many cases chosen to represent natural or typical images that a class of processing techniques would need to deal with. Other test images are chosen because they present a range of challenges to image reconstruction algorithms, such as the reproduction of fine detail and textures, sharp transitions and edges, and uniform regions.

Testing Image Processing: Testing has always been part of Software Carpentry, but it's also always been one of our weak spots. We explain that testing can't possibly uncover all the mistakes in a piece of software, but is useful anyway, then talk about unit testing and test-driven development. Separately, in the extended program design example, we demonstrate how to refractor code to make it more testable.

What we don't do is show people how to test the science-y bits of scientific software. More specifically, our current material doesn't contain a single example showing how to check the correctness of something that does floating-point. You won't find much mention of this in books and articles aimed at mainstream programmers either: most just say, "Oh, round-off," then tell you to use an almost Equals assertion with a tolerance, without telling you how to decide what the tolerance should be, or what to do when your result is a vector or matrix rather than a single scalar value[6][10].

I'd like to fix this, but there's a constraint: whatever examples we use must be comprehensible to everyone we're trying to reach. That rules out anything that depends on knowing how gamma functions are supposed to behave, or what approximations can be used to give upper and lower bounds on advection in fluids with high Reynolds numbers. What might work is simple image processing[4]:

1. It's easy to see what's going on (though using this for our examples does create even higher barriers for the visually impaired).
2. There are a lot of simple algorithms to test that can go wrong in interesting, plausible ways.
3. We're planning to shift our intro to Python to be media-based anyway (using Matt Davis's ipythonblocks and Mike Hansen's novice sub module for scikit-image).
4. People can learn something useful while they're learning about testing.

IV. CONCLUSION

Detection of digital forgery without assistance of signature or watermarking is an emerging topic. We propose a passive approach to detect digital forgeries by checking image quality inconsistencies based on blocking artifacts caused by JPEG compression.

Further work could be done on discovery of other image quality inconsistency measure. Today’s technology allows digital media to be altered and manipulated in ways that were simply impossible 20 years ago.

As we continue to develop techniques for exposing photographic frauds, new techniques will be developed to make better fakes that are harder to detect. The field of image forensics, however, has made and will continue to make it harder and more time-consuming (but never impossible) to create a forgery that cannot be detected.

V. REFERENCES

[1] Hany Farid, ”Image Forgery Detection”, IEEE SIGNAL PROCESSING MAGAZINE, MARCH 2009, pp. 16-25.

[2] M. K. Johnson and H. Farid, “Exposing digital forgeries in complex lighting Environments,” IEEE Trans. Inform. Forensics Security, vol. 3, no. 2, pp. 450– 461, 2007.

[3] 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.

[4] A. Swaminathan, M. Wu, and K. J. R. Liu, “Digital image forensics via intrinsic fingerprints,” IEEE Trans. Inform. Forensics Security, vol. 3, no. 1, pp. 101–117, 2008.

[5] S. Ye, Q. Sun, and E. C. Chang, “Detecting digital image forgeries by measuring inconsistencies of blocking artifact,” in Proc. IEEE Int. Conf. Multimedia and Expo, Beijing, China, 2007, pp. 12–15.

[6] 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.

[7] A. C. Popescu and H. Farid, “Exposing digital forgeries in color filter array “Interpolated images,” IEEE Trans. Signal Processing, vol. 53, no. 10, pp. 3948–
3959, 2005.

[8] 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.

[9] B. Mahdian and S. Saic, “Blind authentication using periodic properties of “interpolation,” IEEE Trans. Inform. Forensics Security, vol. 3, no. 3, pp. 529–538,

[10] 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.