Statistical Approach for Copy Move Forgery Detection

Authors:

Neetu Yadav
M. Tech. (Information and Network Security) Student
Institute of Technology,
Nirma University,
Ahmedabad.

Rupal Agravat
PhD. Scholar,
Institute of Engineering and Technology,
Ahmedabad University,
Ahmedabad.

ABSTRACT

With the range of photo manipulation tools now available nearly anyone can modify and change an image's interpretation by vast degree. An image can be called as a chronicle of visual perception. Copying and pasting a patch of an image on to other part in the same image is the main essence of a copy- move forgery and can be employed for many malicious purposes.

Malicious image manipulations are inimical as they can lead to serious changes to the information that is perceived by the human mind. Many techniques to detect copy-move image forgery using feature descriptors have been used in the past. SIFT features are considered as a robust scale, rotation, translation and affine invariant feature. We have used the approach of clustering similar SIFT feature descriptors and propose to extend the copy move region detection by introduction of segmentation mechanism for precise detection of the forged region.

Index Terms: Copy-move, image forgery, image forensics, SIFT features, statistical methods, PCA, SVD, eigen vectors.

I) INTRODUCTION

Image forensics is a sub branch of multimedia forensics wherein still and digital images are analyzed. Image forensics or photo forensics is a burgeoning field with many researchers aiming for better reliability on the originality and authenticity of different aspects of the digital image or photographs. The objective of most image forgery detection is to localize and report the tampered region of the image from a given digital test image. Image tamper detection on the other hand aims to just establish the fact whether if an image has been forged or not.

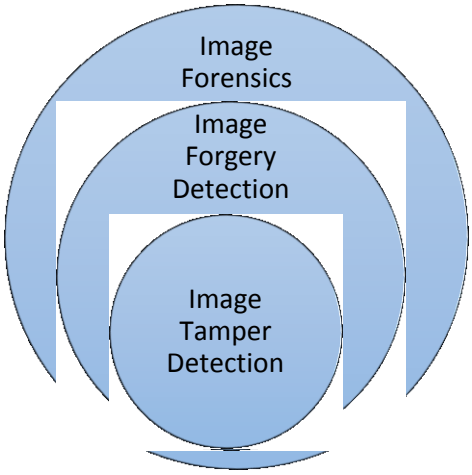

Fig.1:Disintegrating Image Forensics

Image forensics can be considered to include image tamper detection and image forgery localization (Jing Dong Wei Wang & Tieniu Tan, 2009). The applications of image forensics range across widely different domains like Criminal Investigation, Insurance Processing, Photojournalism, Digital Evidence Tampering, Scientific Research, Political Propaganda, etc. Exponential increase in circulation of digital images in the internet and social media has lead to pervasiveness of images with no authenticity.

Most common and notorious image manipulation techniques can be termed as image splicing and copy move forgeries. They are often used to manipulate or hide or add the details present in the original image. Image forgery can be divided into three broad groups Image Splicing, image retouching and copy-paste forgery.

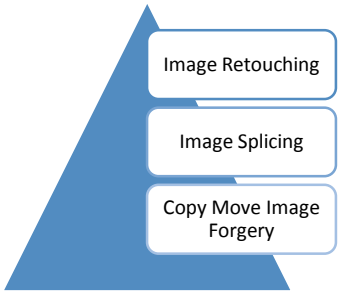

Fig.2: Image Forgery Types

In image splicing a part of image is taken and pasted onto a different image. In copy move forgery a part of image is taken and pasted into the same image consequently hiding some detail or including some extra details. In image retouching the goal is to perform targeted image manipulations on very small regions in the image. Often used in advertisement industry to beautify the face or reduce excess fat, etc. Of the above all forgeries, copy move is the most easiest performed and can leave no visual clue even when done by a layman.

Image Tamper detection can be divided into active and passive techniques (M.Deriche & M.Ali Qureshi, 2014). Both these techniques work with certain assumptions. For active techniques its assumed that the image has been properly watermarked with signature. For passive techniques its assumed that the image consists of no such watermarks and is a potential forged image.

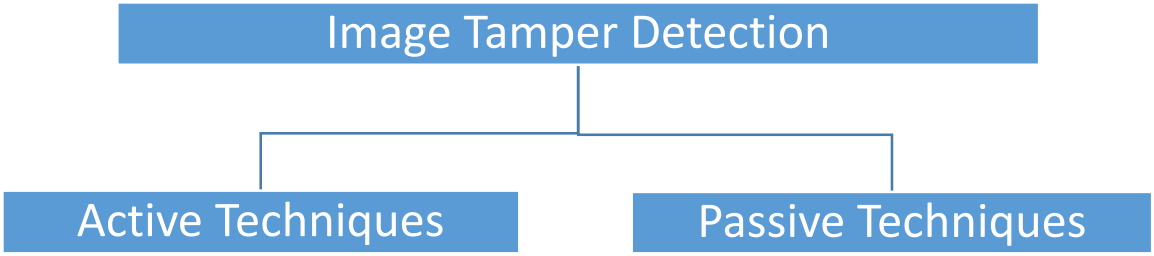

Fig.3: Types of tamper detection Techniques

Active image tamper detection techniques try to retrieve some details from the image to establish the originality and authenticity of the image. For this technique to function they require that images be embedded with a digital signature which may be the device information or the owner information. Such inclusion of details into the image is not an on-board process in most off the consumer based camera devices. Often in criminal investigations images involved seldom contain watermark in them.

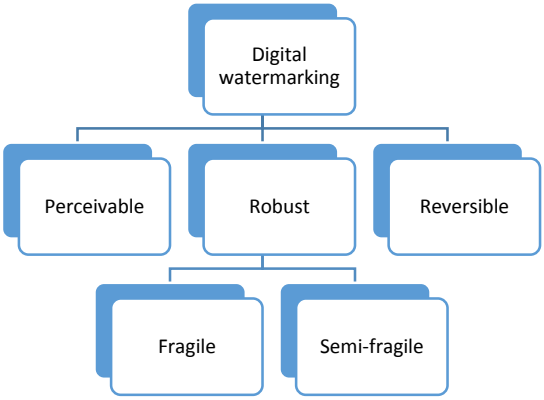

Fig.4: Digital watermarking

Active image forensics uses watermarking technology to embed authentication information into original digital image so that the authentication information can be used for forensic analysis when image has been invaded. Feature of the watermarked images include difference between the original and watermarked image and the watermark is present everywhere. A digital watermark is a kind of marker that is covertly embedded in a noise tolerant signal such as an audio or image data. Its used to identify the ownership of the copyright of image.

Digital watermarks can be further categorized based on the invisibility, robustness and capacity parameters. A fragile watermark is that which fails to be detected in the presence of slightest manipulation. The semi fragile watermark resists initial transformation but fails detection in malignant transformations.

Passive image tamper detection techniques also called as blind detection techniques are more suited for criminal investigations and above stated applications as it can also localize the forged region with maximum accuracy. Passive techniques do not utilize any pre-existing details from the image and analyse the full image. A soft categorization of these techniques is as follows: Source identification, real image vs computer generated image and forgery localization.

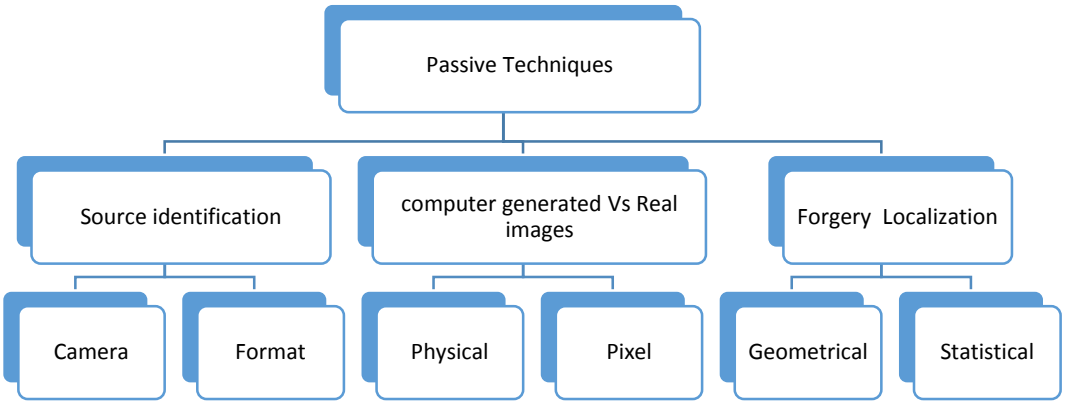

Fig.4: Passive Techniques

Source identification can be carried out by camera and format based forgery detection techniques wherein the source device of the image is identified by examining the noise pattern in the image or the jpeg packaging of the image. This often establishes if an image is tamper with or not, forgery localization seldom performed using above techniques.

Real image Vs Computer generated image are often computer graphics technology is used to create image with striking similarity with real world. Hence it becomes difficult to differentiate the real image form the computer generated image. Image forgery detection techniques like the physical based and pixel based have been to detect apart the real images from computer generated regions.

Forgery localization aims at pure localization of the tampered or forged region in the image. The techniques involved include the Geometric based and the Statistical based techniques. The geometrical based techniques analyse the images by exploiting the fact that measurements of objects in real world and their relative position with respect to the camera utilizing the essence of projection geometry. In statistical based techniques the images are analysed for statistical anomaly present in the tampered images. It is often said that though a forgery may not

leave a visual clue its often possible to detect its presence by performing statistical analysis of the pixel properties with respect to its neighbours.

One or more of the above techniques are often combined to detect forgery in digital images. A combination of Pixel based and statistical techniques have been utilized in the past to detect tamper and also localize the image forgery. **Pixel based techniques** can be divided into keypoint based and block based techniques based in their methodology of operation.

In **block based techniques** the image is divided into fixed size blocks after which pixels of each block are analyzed for presence of collective similarity with respect to other blocks present in the image. This techniques has a variety of drawbacks as it cannot scale well for large sized images and also if the forged region is much smaller than the block size its detection is highly error prone.

In **keypoint based technique**, pixel points in the image having certain characteristic are extracted in the form of features and analyzed with other points present across the image. This comparative analysis exposes the presence of forgery in image. Many keypoint based techniques like SURF, SIFT, RANSAC …etc, exist but their performance varies based on the number of descriptors, points detected etc. This techniques scales well for all size of images.

In **statistical based technique** the underlying image structure of the image is computed using the technique such as the principal component analysis , the eigen vector estimations, dimensionality reduction techniques, the svd computation and computations based on the features that might be extracted using above techniques pave way for the automation of the detection process based on the approximate values. For, example the principal component analysis cannot scale well for large images and also rotated as well as scaled copy move regions not detected.

Principal component analysis technique is used to reduce the large-dimension feature set derived from the text image. Principal component analysis has been developed based on the matrix theory for singular value decomposition.

The reduced dimension representation developed by svd, provides a space in which to identify similar blocks in the presence of corrupting noise, as truncation of the basis will remove minor intensity variations.

II) PREVIOUS WORK

Two of the pioneering researchers in the field of image forensics are Jessica Fridrich and Hany Farid. Some of their current publications for image forgery detection include detection of performing camera based ballistics that is to match the digital image with the source camera (used in image tamper detection) and analyzing the shadow and lighting conditions in the image (used in splicing detection).

Figure.4 lists various techniques that have been in the past to detect the copy-move forgery and has been categorized based on the methods invariance towards Rotation, Scaling and Translation (RST). Further, due to lack of proper forgery detection algorithm evaluation and benchmarking the developed approaches merely present the proof of concept and also the attack techniques for the proposed algorithm are barely studied leading to questioning the robustness of the methodology.

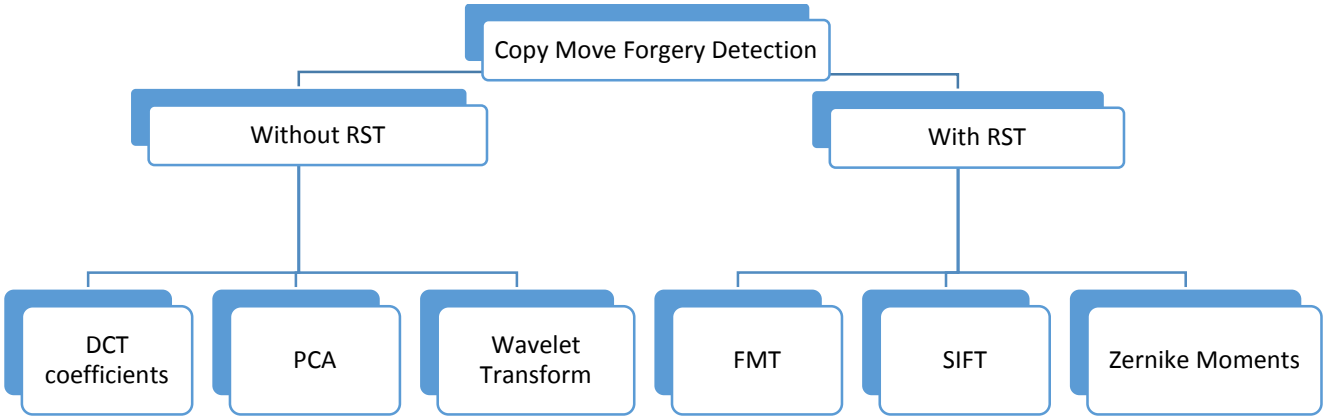

Fig.4:Survey of CMFD Techniques

J. Fridrich, D. Soukal, and J. Lukas [1] in Paper Entitled "Detection of copy-move forgery in digital image" used DCT method where images are divided into blocks and Discrete Cosine Transform performed on each block after which DCT coefficient similarity analysis is performed to check for duplicate regions in the image. The dimension of features analysed here were 64 with forgery localization accuracy of 8X8. But the performance of this technique diminished drastically with noisy images.

A. C. Popescu, and H. Farid [2] in Paper Entitled "Exposing digital forgeries by detecting duplicated image regions" used PCA method where the images are divided into blocks of 8X8 and 32 dimension feature were extracted and from each block and matched for similarity with every other block. The performance of this method depends on high time complexity and is not scalable across different image forgery techniques like scaling and rotation.

Due the intrinsic problem of using block based techniques. Keypoint based techniques were used were certain keypoints were extracted. Two types of matching techniques were proposed the rule based and the robust matching techniques for features extracted from the image.

III) PROPOSED WORK

Our method of copy move forgery detection is based on the idea of SIFT feature similarity present in the copy moved regions as it is easy to implement and robust to various types of copy-move post processing.

In this method first the input image is converted to gray scale then SIFT features are extracted and then clustering operation is applied on them which is followed by the segmentation of the indicated regions from clusters in previous step. Among the regions detected with higher similarity between SIFT descriptors are grouped and the forged regions are located in the image.

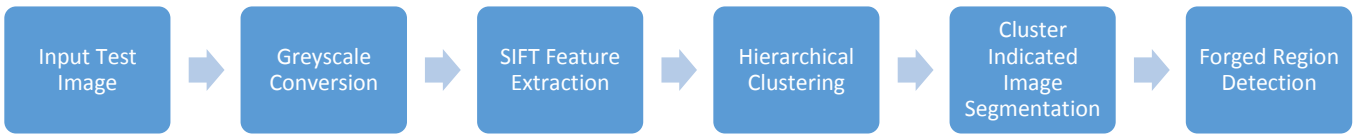

Fig.5: Proposed method

SIFT features are RST (Rotation, Scale and Translation) invariant and hence are good candidates for modified duplicate region detection often present in copy move forgery. Application of clustering operations in the transformed domain helps reduce the domain of comparison and speeds up the operation. Further the segmentation of the image facilitates the detection of forged region chunks in the image.

The proposed algorithm can be used to perform further analysis of forged regions were the region around the detected forgery can be subjected to analysis with respect to surrounding image pixels in the process to recover the original image components.

IV) FINDINGS

Using this method, Images which are forged can easily be detected especially forgery detection can also be done when the objects on which transformations (rotation, scaling, translation) are applied and then moved at different location within same image.

V) CONCLUSION

An progressive method to enable image forgery detection based on SIFT features is proposed. Given a suspected test photo, it can reliably detect for duplicate regions. The presented technique can show effectiveness with respect to multiple

modifications such as scaling, rotation and translation to copy-move forged region in the image. The clustering phase has been extended by way of an image segmentation procedure. SIFT based image forgery detection techniques can further be extended to include the detection of other image forgery apart from copy-move. Implementing the machine learning techniques to automate this process further enhances the applicability and deployment of algorithm for real time forgery detection.

VI) REFERENCES

[1] David Soukal Jessica Fridrich and Jan Lukas. Detection of copy-move forgery in digital images. Digital Forensic Research Workshop, 2003.

[2] A. C. Popescu and Hany Farid. Exposing digital forgeries by detecting duplicated image regions. IEEE Signal Processing Magzine, 2009.

[3] Irene Amerini, Lamberto Ballan, Student Member, Roberto Caldelli, Alberto Del Bimbo, and Giuseppe Serra. A sift-based forensic method for copy move attack detection and transformation recovery. 2011

[4] Gajanan K. Birajdar and Vijay H. Mankar. Digital image forgery detection using passive techniques: A survey. Digital Investigation 10, no.3, 2013.

[5] Lt. Dr. S.Santhosh Baboor B.L.Shivakumar. Detecting copy-move forgery in digital images: A survey and analysis of current methods. Global Journal of Computer Science and Technology.

[6] Sonja Grgic Dijana Tralic, Ivan Zuoancic and Mislav Grgic. Comofod-new database for copy-move forgery detection, 2013.

[7] Hany Farid. Image forgery detection- a survey. IEEE Signal Processing Magzine,2009.

[8] David G.Lowe. Object recognition from local scale-invarient features. 7th, IEEE International Conference ,vol2, 1999.

[9] Chaoqun Ma Guojun Gan and Jianhong Wu. Data Clustering : Theory, Algo-rithms and Applications. 2007.

[10] Wiem Taktak Jean-Luc Dugelay and Judith A. Redi. Digital image forensics: A booklet for beginners. Eurecom, France.

[11] Ji r Matas Karel Lenc and Dmytro Mishkin. A few things one should know about feature extraction, description and matching. 19th , Computer Vision Winter Workshop, 2014.

[12] M.Deriche and M. Ali Qureshi. A review on copy move image forgery detection techniques. Multi-conference on system, signal and devices, IEEE, 2014.

[13] Corneliu Florea Mihai Ciuc Peter Corcoran, Cosmin Stan and Petronel Bigioi. The good, the bad, and the (not-so) ugly, 2014.

[14] Jing Dong Wei Wang and Tieniu Tan. A survey of passive image tampering detection. International Conference on Pattern Recognition, 2009.