A Survey on Approaches for Detecting Forged Image

Mrs. Rutuja G. Tendulkar
Student of M.E (I.T)
rutu_tendulkar@yahoo.co.in
Department of Information Technology,
V.E.S.I.T, Mumbai, India

Mr. Manoj Sabnis
Associate Professor,
manojsab67@yahoo.co.in
Department of Information Technology,
V.E.S.I.T, Mumbai, India.

***ABSTRACT*:** Digital image forgery is the process of changing contents of a document and representing the changes as true copies of the original. Due to the availability of powerful image processing and editing softwares, it is easy to manipulate and edit digital images. Authenticity of an image received on communication channel is important now a day as digital images are accepted as evidence into court of law. This paper includes active and passive approaches used for finding forged image. These approaches cover three techniques such as watermarking, hashing and jpeg compression to find manipulated image.

Keywords*: image forgery, watermarking, hashing, jpeg compression, quantization table*

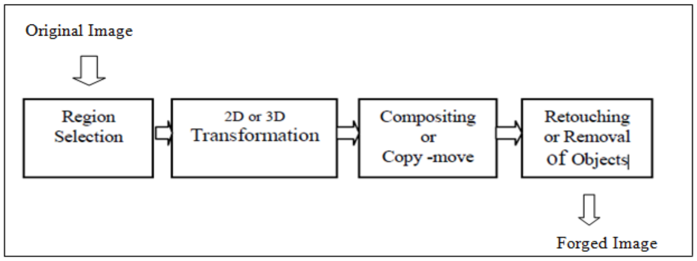
I. INTRODUCTION

A digital image is any photograph, agreement, letter or other written material which may be a scanned or faxed or taken by any digital device. Any modification, alteration or enhancement of the digital image after it leaves the camera is treated as image forgery. Digital images play important roles in many fields such as medical, journalism, scientific publication; digital forensic etc. The digitally forged images are sometimes so real and it cannot be distinguishable from the original image. Technology advancement allows people to easily alter the content of digital multimedia. Manipulation to digital images is done for hiding some meaningful or useful information, to create misleading images or to make forged images. This brings a new challenge toward establishing trustworthiness of digital multimedia i.e. authentication of the image received in a communication.

(a)Process of forgery creation:-

The image forgery creation process involves selection, transformation, composition of the image portions and retouching of the final image [1]. In this process, forger first selects the complete image or specific region of image for performing forgery. Then forger transforms the image into frequency domain. From this frequency domain some portion of image is transformed into another image using different forgery techniques. Finally the forged image is retouched to remove the remaining objects.

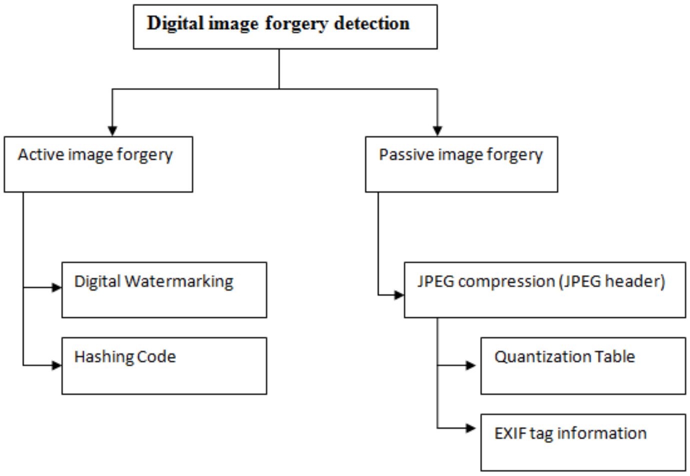
Block diagram [1]:-



(b)Types of Image Forgery creation: It includes enhancing, composition and copy-move forgery [2][3].

- **Enhancing**: - It is also known as image retouching. It is less harmful kind of digital image forgery. It does not significantly change an image but enhances or reduces certain features of an image. One can enhance certain features of an image to make it more attractive but they are ethically wrong.
- **Composition**: - It is also known as image splicing. It is a technique that involves a composite of two or more images which are combined to create a fake image.
- **Copy move forgery**: - It is a technique that copy background or other features from one part to hide or alter other areas of the original. In this type, instead of having an external image as the source, it uses porttion of the original base image as its source. That means source and destination of the modified `image originated from the same image.Part of the original image is copied and moved to desired location and pasted.

II. APPROACHES

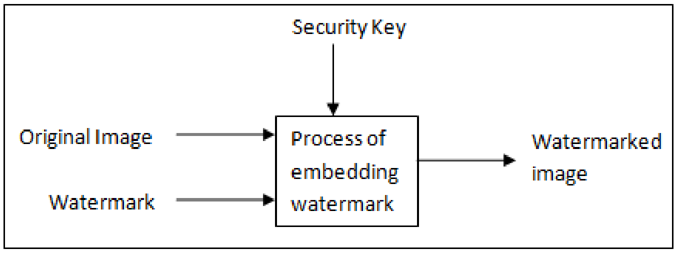


Active approach: - It requires prior information about the image such as watermark embedded inside the original digital image or hashing code computed for the original digital image.

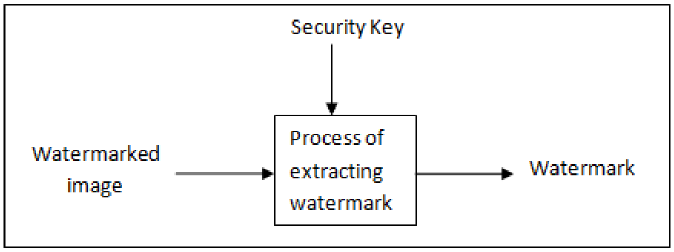
(a) Digital watermarking:-

The initial technique to detect digital image forgery is digital watermarking. The process of embedding additional data along with the digital images is called as digital watermarking [4][5]. i.e. the watermark is inserted into the image, and while detecting forged image, watermark is extracted from received image. Then watermark of original image and watermark extracted from the received image is compared to find manipulation in the image. Any damage into the watermark indicates an image is forged.

Block diagram of a general watermark embedding system:-



Block diagram of a general watermark extraction system:-

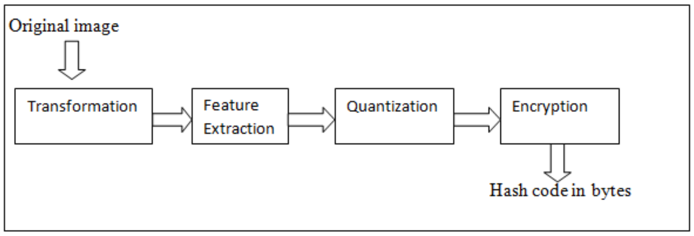


In the watermark embedding system, a watermark is inserted into the original image. A watermark is a binary bit sequence. Security key is used to prevent unauthorized users from generating legitimate watermarked signals. The watermark selected for embedding should not distort the original image transmitted over the network.

(b) Hashing Code:-

An image hash is a distinctive signature which represents the visual content of the image in a compact way i.e. usually just few bytes [6]. Image hash calculated for an image should be robust against any manipulation taking place in an image. Hash code should be unique for each image i.e. it should be different for different images and tampered images. Image hashing technique is used to find the originality of an image received on communication network. An image hash is calculated based on the visual contents of the image with selected features. Then it is sent with the original image in the form of tag. When an image is received on the network, its hash is computed at the destination to verify the reliability of received image. Then newly computed hash is compared with the hash code of the original image. If both differ then it indicates forged image.

Block diagram for Generating hash code for digital image:-



- Original image or the received image is selected for calculating hash code.
- This image is then transformed with the help of spatial transformation such as color transformation where color image is converted into grayscale image to reduce computational cost of feature extraction. Transformation is done to extract features depending upon values of image pixels or frequency coefficients.
- In the next stage of feature extraction, some features which are essential are selected from the transformed image to generate feature vector. One can select features from the image with specific patterns including edges, corners, salient regions, etc.
- In the quantization stage, quantization is applied to each component of hash vector. It presents the hash in bytes.
- In this stage, short hash of fixed size of 1 byte is encrypted using cryptographic algorithms. This stage includes cryptographic hash functions i.e. SHA to generate the final hash of fixed size.

Once the hash is calculated it is send with the image. To find forged image we compare the hash values that are generated in the sender and receiver side. Receiver generates the hash value for received image and compares it with hash value of sender's image. If both matches then image is not forged. If both do not match means the image is forged.

Passive approach:-

It does not require explicit prior information about the image [1]. This approach deals with source identification and forgery detection. Source identification means finding the source digital device which is used to capture the image such as mobile phones, cameras, etc.These devices are used to discover evidence of tampering by assessing the authenticity of the digital media by recovering information about their history [1].

Jpeg compression:-

Most of the digital cameras can produce images in JPEG file format.JPEG file format is a source of data that can be used for the purpose of detecting forged images.

In jpeg compression an image travels through various processes such as dividing an image into multiple blocks (N×N), then transforming each block with DCT transform into frequency domain values i.e. AC and DC coefficients, then quantizing DCT values to reduce the number of bits required to store the values and then encoding the values with Huffman coding. In this process of compression, JPEG file format records the information about the image in meta-tags. The information includes quantization matrix.huffman table, make and model of source device, time of creation and other parameters. The contents to be stored in tags, sequence of tags and available tags depends on the image as well as the digital device used to capture it and a software with which it is modified. Every image file contains attributes which can be used for detecting forgery when we do not have original image.

For detecting forged image we can also use camera attributes. It includes cameras of different make, model, resolution, and quality. Cameras often support multiple resolutions and qualities, each of which yields images with different JPEG compression parameters. Every digital device when it save the image, compress it with its own set of quantization tables. Quantization table of each digital device and editing software is unique.

JPEG header format: -

The JPEG format headers can be exploited for image authentication. A camera signature is extracted from a JPEG image consisting of information about quantization tables, Huffman codes, thumbnails, and exchangeable image file format (EXIF) [1].

JPEG header includes Image Parameters, Thumbnail parameters and EXIF metadata parameters which can be used for detecting forged image [1][7].

Image Parameters:-

It contains camera signature indicating the image dimensions, quantization table, and Huffman code. Each camera has its own sensor resolution so image dimension can be used to differentiate between cameras with different sensor resolution. It contains 8×8 quantization table presented with one dimensional array. Each component has a separate quantization table, so three tables are maintained for luminance(Y), chrominance (Cb) and Chrominance (Cr) respectively. All three tables give total 192 values which can be extracted from header. The Huffman code is specified as six sets of 15 values corresponding to the number of codes of length 1, 2, 15: Each channel is represented with two codes where one code is used for ac coefficient and the other is used for dc coefficients. In total, we extract 284 values from the full resolution image: 2 image dimensions, 192 quantization values, and 90 Huffman codes.

Thumbnail Parameters:-

A thumbnail image is a smaller version of full resolution image. In jpeg compression process, a thumbnail version is often embedded in the JPEG header. The thumbnail is then compressed and stored in the header as a JPEG image. From the thumbnail image, we can extract image dimension, quantization table and Huffman code. As with the full resolution image, we have found that the chrominance channels are encoded with the same parameters. Some camera manufacturers do not support thumbnail image embedding in header. So if thumbnail image is not available then thumbnail parameters are assigned a zero value. In total, we extract 284 values from the thumbnail image same as image parameters: 2 thumbnail dimensions, 192 quantization values, and 90 Huffman codes.

EXIF Metadata Parameters:-

Camera signature contains EXIF metadata parameters which can be extracted for authentication of an image. EXIF parameters include information about an image as well as the source digital device used to capture the image. The information includes make and model of image description, image length, image width, source device, file size, date/time of creation, x-resolution, y-resolution, resolution unit, and etc.Listing of attributes of an image depends on the source device or editing software.

Image Authentication:-

We extract 284 header values from the full resolution image, a similar 284 header values from the thumbnail image, and another set of attributes from the EXIF metadata. These values form the signature can be used to find forged image. Specifically, the signature and camera make and model are extracted from the EXIF metadata and compared against authentic image signatures extracted from the same camera make and model.

EXIF tags of this image are a clear indication of image manipulation. The Software tag displays software used for editing the image, and the original date and time does not match last modification date and time.

Another method to find forged image is to find quantization table from received image. Every image when created with digital device it stores its device type as source in the attribute. When a device compress an image it uses quantization table of its own whereas any editing software that is used for altering an image uses its own quantization table when it does compression. Then the altered image contains the quantization table of its editing software. We can extract make and model from an image and compare its quantization table with the quantization table of received image. If both are same then received image is not forged.

An image's EXIF metadata can be relatively easily edited to alter the camera make and model [7].In this situation we can extract the quantization table from the received image and then compare it with quantization table all digital devices. If it matches to any one of them then the received image is not forged.

III. Result and Discussion

(a) Digital watermarking is a technique to authenticate the digital image received on communication channel. This technique embeds some authorized signatures, referred as watermark signals, into original digital images invisibly or visibly depending on the application scenarios. At the receiver side, the watermark can be extracted and compared with original watermark.

Drawback of watermarking:-

- Watermark must be inserted either at the time of recording the image, or later by a person authorized to do so. This limitation requires specially equipped cameras or subsequent processing of the original image.
- Inserting a watermark sometimes may result in distorted image.
- Watermark may vanish if someone manipulates the images.
- Resizing, compressing images from one file to another may diminish the watermark and it becomes unreadable.

(b) Computing Hash code for received image and comparing it with original image's hash code helps in finding forged image. If both the hash codes are different then the received image is forged.

Drawback of Hashing:-

- Computing and sending the hash value requires special calculations and time.
- Hash value send with the image can be extracted by manipulator to generate the same hash for forged image i.e. it is less secured.
- Extracted features may not be the ideal points.
- Accuracy of finding forged image is less.

(c) Jpeg compression:-With JPEG header information it is easy to extract Quantization table used for compressing digital image. Comparing extracted quantization table with quantization table of source make and model of the image helps in finding forged image. If there is difference in values of quantization tables then the received Image is forged.

Drawbacks of Jpeg compression:-

- An image's EXIF metadata can be relatively easily edited.
- It requires extra memory space for image JPEG header.
- This technique requires camera signatures from a wide variety of cameras and cell phones. This poses significant challenges as new cameras and cell phones are constantly released.

IV. Conclusion

Watermark image can be used for detecting a forged image by examining embedded watermark inside it. But watermark distorts the image appearance so to improve image quality we can use hash code for finding forged image. Hash code is computed with extracting features from image and sends with the original image. By computing hash code for received image we can compare hash code of both and find forged image. But hash code computation requires more time and always the extracted features are not ideal every time, so we can use JPEG header to find forged image. With jpeg header we can compare values extracted from header with camera signature from which an original image is taken. But the problem with this method is that it works only with jpeg images.

Watermarking is used to identify enhancing and copy-move forgery. Hash code and JPEG compression both support detection of image forgery of type enhancing, composition and copy-move.

Watermarking technique and hashing technique both require knowledge of original image whereas in JPEG compression technique without original image we can find forged image.

V. References

1. Anupama Saini published a paper titled ***'A Survey on Passive-Blind Image Forensics'*** in 2nd International Conference on Role of Technology in Nation Building (ICRTNB-2013) ISBN: 97881925922-1-3.

2. S. Devi Mahalakshmi (Assistant Professor) 1, Dr. K. Vijayalakshmi (Professor) 2,E. Agnes (PG Scholar) published a paper titled '***A Forensic Method for Detecting Image Forgery***' in 2013 IEEE International Conference on Emerging Trends in Computing, Communication and Nanotechnology (ICECCN 2013).

3. Granty Regina Elwin J #1, Aditya T S #2, Madhu Shankar S #3 published a paper titled '***Survey on Passive Methods of Image Tampering Detection***' in International Conference on Communication and Computational Intelligence – 2010

4. B Manoj Kumar,DR.T.V.S Gireendranath published a paper titled '***Novel Invisible And Blind Watermarking Scheme For Copy Right Protection Of Digital Images'*** in International Journal of Engineering Research & Technology (IJERT)Vol. 1 Issue 7, September – 2012,ISSN: 2278-0181

5. Prabhishek Singh, R S Chadha published a paper titled ***'A Survey of Digital Watermarking Techniques, Applications and Attacks'*** in International Journal of Engineering and Innovative Technology (IJEIT) Volume 2, Issue 9, March 2013

6. S.Battiato,G.M.Farinella,E.Messina,G.Puglisi published a paper titled '***Understanding Geometric Manipulations of images through BOVW-Based Hashing***' in 2011 IEEE

7. Eric Kee, Micah K. Johnson, and Hany Farid published a paper titled '***Digital Image Authentication from JPEG Headers***' in IEEE Transactions on information forensics and security', VOL. 6, NO. 3, September 2011