Research on tagging of Copy-Move bogus and ordinary imagery by ORB Features and SVM Classifier

Ankush Kapoor[1], Khusboo Khanna[2]
[1]Research Scholar - M.Tech (CSE), [2]A.P. (CSE)
CGC Technical Campus, Jhanjeri, Mohali[1]

Abstract— The input images are divided into overlapping and regular image blocks by the use of existing block-based forgery detection methods, and then image pixels or transform coefficients matched block by which the tampered regions are obtained; and the keypoint-based forgery detection methods in which the image keypoints are extracted and match them for the duplicated regions identification. In this forgery detection method which divides the input image into over-lapping rectangular blocks, from which matches the blocks of the quantized Discrete Cosine Transform (DCT) coefficients for the tampered regions finding. The Principal Component Analysis (PCA) is applied for the reduction of the feature dimensions. In the RGB color components, the direction information as block features is used. The Discrete Wavelet Transform (DWT) and Singular Value Decomposition (SVD) used for the extraction of the image features. Some limitations are there in the existing systems, although in forgery detection, effective are these scheme. Thus, dividing the host image into over-lapping rectangular blocks, computationally this would be expensive as the image size increases. The forgery regions geometrical transformations cannot be significantly addressed by the methods. They have low recall rate because of the regular shape of their blocking method.

Keywords - DCT, SVD, DWT, PCA

I. INTRODUCTION - DIGITAL FORGERIES DETECTION NEED

The powerful programs availability of digital image processing, such as Photoshop, makes the digital forgeries creation relatively from one or multiple images. As shown by the newspaper cutout, for the creation of the composite image uses three different photographs: The White House, Bill Clinton, and Saddam Hussein images. The White House was rescaled and blurred for an out-of-focus background illusion creation. Then, two different images cuts off the Bill Clinton and Saddam and on the image of the White House, it is pasted. The fact that sophisticated tools are used for digitally manipulating the images and video to create threatening non-existing situations due to which the credibility and value of presented video tapes and images is diminished in court as evidence independently of the fact either the video is in a digital or analog form. The analog video stream is digitized easily for an analogue video tampering, and uploading it into a computer, forgery is performed, and then the result is saved on an ordinary videotape in the NTSC format. As expected, worse the situation will get as the needed tools by which the forgeries are performed moving from research labs to commercial software.

Despite the fact that the research community recognizes the digital forgeries detection need, and currently very few publications are available. The proposed digital watermarks as a means for fragile authentication, content authentication, detection of tampering, localization of changes, and the original content recovery[1].

A. Copy Move Forgery

Nowadays a variety of applications rely on digital images. These include newspapers, tabloid magazines, scientific Journals, fashion industries, court halls and many others. Today, almost everybody can record, store and share a large amount of digital images because of the spread of easy and cost effective device that enables the acquisition of visual data (Shiva kumar and Baboo, 2011). At the same time, image editing software is widely available which makes it extremely simple to manipulate the content of the image. This can be achieved through creating new images by tampering and counterfeiting the visual content in an expert – like method. Current software allows users to create computer graphics that can't be distinguished from real photos or even to generate hybrid generated visual content (Meyer, et al., 1986). Such developments lead us to ask different forensic – related questions.

B. Digital Image Forensics

Digital forensic science is a modern branch of science which aims to reconstruct events and identify entities involved. This field deals with the study of deciding the originality and reliability of digital media such as images. Strictly speaking, the term forensic implies the application of scientific approach on the investigation and detection of a crime. Such forensic analysis serves in providing proof or evidence at courts. Recently, Digital images have widely spread leading to the use of digital image forensic in broader context of situation (Kirchner, 2012).

C. *Image Source Identification*

Image Source Identification is considered one of the basic problems that the techniques of digital image forensic try to solve. This technique aims in particular to recover the used type of imaging devices such as digital cameras, scanners, mobiles, computer, and graphic technologies. By focusing on visual data, the specific model or brand of such devices can be identified, since each digital imaginary device leaves a unique imprint on the acquired image during the processing operation or the storing phases. This difference arises from the fact that images have different characteristics because of using different physical devices and different image processing techniques.

D. Digital Image Forgery Types

The digital image semantic contents may be altered through information removal from that image, or extra information is added to it. There are numerous ways that forgers may use to achieve that. In addition, different criteria are used for classifying those techniques such as the involved number of images in the manipulation operations (Al-Hammadi, 2013). ulong, 2014).

II. LITERATURE SURVEY

Gomase and Wankhade (2014) In this paper, a technique that is proposed in which intensity of the local changes in the image is found out by applying DWT. For the removal of noise, a median filter is applied. For detection process, dividing the image into overlapping blocks, then storing in a matrix and finally sorting the matrix. Finally, using the matrix by which the copy-move regions are located through pixel matching. This method is useful only when preprocessed are the images, but only copied regions shifting.

- **Hashmi, et al.(2014)** presented a methodology using (DyWT) and (SIFT), which ensures better detection rates after preprocessing and other attacks. First the image is converted to wavelet form (DyWT) for its decomposition into four parts: LL, LH, HL, and HH .SIFT then applied to LL part that most of the information is contained, to obtain the multispectral components and feature vector descriptors.
- **Zhao and Guo (2013)** This paper proposes a robust method for detection of copy-move forgery which is based on applied to each block. The robust representation is obtained by the quantization of the DCT coefficients which is followed by the quantization blocks division into non-overlapping sub-blocks. On each sub-block applying SVD.
- **Al-Sawadi et al.(2013)** This paper presented a copy-move image forgery detection method which based on Local Binary Pattern (LBP) and neighborhood clustering. In the proposed method, three color components are what image is decomposed first in. Each component overlapping blocks is used to calculate the LBP histograms. Then calculating the histogram distance between the blocks and retaining the minimal distance in the block-pairs. If the retained block-pairs in all three color components are present and as primary candidates, they are selected.
- **Davarzani et al.(2013)** In this paper, the method proposed is a tampering detection method which is based on LBP. The copied regions is detected by this algorithm even if the forged region geometry is polluted further by noise, blurring, JPEG compression, scaling or rotation in multiples of 90-degree. In this algorithm the image translation is basically into gray scale and then subdivision is into overlapping blocks. For each block of multi-resolution Local Binary Pattern (MLBP) features are identified by applying the LBP operators of different type.
- **Zhong and Xu (2013)** This paper presented mixed moments based method. First, uses the Gaussian pyramid transform for the extraction from the image of the low-frequency information then divided it into overlapping blocks.
- **Muhammad (2013)** In this paper, a multi-scale local texture descriptor is proposed for image forgery detection. (Hussain et al., 2012) presented the same approach, but with the applying difference of the un-decimated wavelet transform for the channel extracting the lower sub-band. Before doing that, the chromatic channel is into which it is decomposed on an input image.
- **Amerini et al.(2013)** In this paper, for feature extraction a Scale Invariant Feature Transform (SIFT) is employed, the J-Linkage algorithm bases localization combined for detecting tampering. For the image extracting the SIFT features. The g2NN algorithm is used for the matching of the feature vectors afterwards.
- **Cao et al.(2012)** This paper presents region of duplication detection algorithm on which improved DCT depends upon and low computational complexity exhibited. This method and the other DCT-based methods profound difference is that here characterizing the quantized block by a circle block.
- **Shao et al.(2012)** This paper proposed an algorithm which computationally is a complex copy -move forgery detection algorithm. The circular window expansion and phase correlation is on which this algorithm depends upon. The circular window scans the image which is expanded then into a normalized rectangular block with the use of bi-linear interpolation.
- **Muhammad, et al.(2012)** In this paper, a method is proposed using un-decimated Dyadic Wavelet Transform (DyWT), in which shifting invariance is property in it and thus for data analysis this is more suitable than Discrete Wavelet Transform (DWT). First, approximation (LL1) and detail (HH1) sub-bands into which the input image is composed of. Then the division of the LL1 and HH1 sub-bands into overlapping blocks and calculating the similarity between those blocks.
- **Hussain et al.(2012)** a multi resolution Weber local descriptor (WLD) system is proposed in this paper, in which "Weber" law is used so as highly textured images is detected with the transformations and shapes of different types of the copied regions. Firstly, changing the colored image into YCbCr color mode in which the color components is stored in chrominance and luminance factors that gives more information than can be done by the human eyes.
- **Quan and Zhang (2012)** In this paper, a texture based copy-move forgery detection scheme is proposed in

digital images. Firstly, using the intrinsic dimension estimation method for the segmenting of the image and then identifying the copy-move forgeries in the image regions in the same texture presence.

- **Hsu and Wang (2012)** In this paper, a new forgery detection system is proposed which is based on Gabor filter. Considering the Gabor filter with different scaling factors, rotation angles and frequencies for generating an image Gabor feature representation. For two images comparisons, applying their Gabor features for finding if in them there is any similarity.
- **Jing and Shao (2012)** In this paper, a copy-move forgery detecting method is proposed which is based on local invariant feature matching. With the feature points matching, the copied and pasted regions are located by this method. The SIFT algorithm uses local features extraction and feature points to be detected. The k-d tree and Best-Bin-First method based is the matching local features.
- **B.L.Shivakumar**[10] **(2012),**In this paper, a technique is proposed for the detection of Copy-Move Forgery which is based on SURF and KD-Tree for matching the multidimensional data. Their method demonstrates with images of high resolution affected by copy-move forgery.
- **Li Jing et. Al.** [2] **(2012)** In this paper, firstly analyzing and summarizing block matching technique is proposed, then a copy-move forgery detecting method is introduced which is based on matching of local invariant feature. The copied and pasted regions are located by it by feature points matching. Then, feature points are detected and local feature is extracted by using Scale Invariant Transform algorithm.
- **Xunyu Pan et. al** [6] **(2011)** In this paper, a method is suggested for the detection of duplicated regions suggestion with continuous rotation regions. The new method was based on the image SIFT features.
- **Yanjun Cao, Tiegang Gao** [12] **(2011)**, This paper presents an efficient and robust approach for the detection such specific artifact. Firstly, the division of original image into fixed-size blocks, and discrete cosine transform (DCT) to each block is applied, thus, each block represented the DCT coefficients. Secondly, block transforming each cosine is represented by a circle block and extracting the four features for reducing the each block dimension.
- **Preeti Yadav, Yogesh Rathore**[8] **(2010)**, In this paper, an improved algorithm is proposed which is based on Discrete Wavelet Transform (DWT) used for the detection such cloning forgery. In this technique, applying DWT (Discrete Wavelet Transform) to the input image for yielding a dimensional representation reduction. After that dividing the compressed image into overlapping blocks.
- **Chun Wang et.al**. [9] **(2010)** For detection of the copy-move forgery is situation that is more challenging is for detecting the duplicated region that rotates in some angle before pasting it. The presented method is for detection in limited rotation angles in duplicated regions.
- **Bayram, et al.(2009)** This paper conducts a study for copy-move forgery detection by the use of Fourier-Mellin Transform (FMT). FMT is chosen by them because this method is robust to lossy JPEG compression, blurring, noise, scaling and as post-processing applied by translation effects. At the beginning, dividing the image into small sized several blocks and calculating each block Fourier Transform and ensuring invariant translation transformed.
- **Frank Y. Shih et. Al.,[**7] **(2009)**, In this paper, the techniques discussed is copy-cover image forgery and four detection methods are compared for copy-cover forgery detection, which are PCA, DCT, spatial domain, and statistical domain based. Their effectiveness and sensitivity is investigated under the influences of Gaussian blurring and lossy JPEG compressions.
- **Bayram et. al** [11] **(2009)** This paper proposed a method Fourier Mellin Transform (FMT) which is applied on the image block. Firstly, the Fourier transform representation is obtained by them in each block, the resulting magnitude values is re-sampled into log-polar coordinates.
- **Fridrich et al.** [4] **(2007)** This paper suggest the the copy-move forgery detection method for detection. In their method, firstly segmenting the image into overlapping small blocks as followed by the extraction of feature.
- **Popescu and Farid (2004)** suggested a method using Principal Component Analysis (PCA). In this method the image is transformed into grayscale and separated into many parts represented into vectors. These parts or blocks are organized lexicographically and PCA is used to represent the dissimilar blocks in a substitute mode. It is proficient for detecting even minor variations resulting from noise or wasted compression.
- **Vincent Christlein** [3] **(2004)** In this paper, the aim is to answer the best performing copy-move forgery detection algorithms and processing steps(e. g. , matching, filtering, outlier detection, affine transformation estimation) in various post processing scenarios.
- **Fredrich, et al.(2003)** In this paper, a method proposed is for the detection of copy- move forgery. The image block using the Discrete Cosine Transform (DCT) and considering their lexicographical sorting for avoiding the computational burden. Once sorted, considering the adjacent identical blocks pair to be copy-moved blocks. This method drawback is that small duplicate regions is not detected.

III. METHODOLOGY

Forgery detection methods become much more complicated to deal with the latest forgery techniques. The forgery detection method divide the input images into overlapping and regular image blocks certain processes needs to be applied for better and efficient results. The methodology of the process is described below-

A. Methodology

Forgery detection methods become much more complicated to deal with the latest forgery techniques. The forgery detection method divide the input images into overlapping and regular image blocks certain processes needs to be applied for better and efficient results. The methodology of the process is described below :

1. Input the different types of images.
2. Extract different type of features.
3. Normalize the features by scaling method.
4. Matching using ORB features.
5. Oriented the feature FAST &amp.
6. Rotated the feature BRIEF.
7. Classification by reducing the false positive error.
8. Post processing by analysis precision.
9. And post processing by recall the precision.
10. Then check the accuracy of the precision.

B. Tool Used

MATLAB is a platform named as matrix laboratory which provide the numerical computing in multi-pattern. It is also called as the 4th generation language of programming. The researcher of Math Works Inc. Dr. Cleve Moler had collected the first form of MATLAB in 1970's. It is developed for the students so that they can access the LINPACK and EISPACK projects without any need of learning the FORTRAN language. By utilizing MATLAB, exploitation of matrix become easy, the in sequence and functions can be plotted easily, calculations can be executed, an algorithm can be implemented, unusual client interfaces can be made additionally interfacing should be probable with the projects which are actualized in varied programming languages like JAVA, C++, C and Python. It is used for matrix theory, linear algebra and numerical analysis. The MATLAB application is fabricated around the MATLAB scripting languages. It also supported, the object oriented programming which combine classes, inheritance, packages, pass-by-illusion semantics. It ropes to develop the applications with GUI (graphical user interface). It likewise has firmly integrated diagram plotting apparatus. Several technical and computing problems can solve by using MATLAB. Lots of researchers used MATLAB for their replication. MATLAB is also especially useful in the field of Sensor network where it consists inbuilt sensor signal, routing Simulink and various tool boxes like signal processing toolbox, sensor network etc. MATLAB is a high-performance language for technical computing. It integrates computation, apparition, and programming in an easy-to-use background where problems and solutions are articulated in well-known mathematical notation. Typical uses include: Algorithm development Modeling, simulation, and prototyping Data examine, exploration, and visualization Scientific and engineering graphics Application development, including Graphical User Interface building MATLAB is an interactive system whose basic data constituent is an array that does not necessitate dimensioning. to write a program in a scalar non interactive language such as C or Fortran.

The name MATLAB stands for matrix laboratory. MATLAB was initially written to make available easy access to matrix software developed by the LINPACK and EISPACK projects, which mutually represent the state-of-the-art in software for matrix computation. MATLAB has evolved over a period of years with input from various users. In university environments, it is the standard instructional tool for introductory and advanced courses in mathematics, engineering, and science. In industry, MATLAB is the tool of choice for high-productivity research, development, and examine.

IV. EXPERIMENTAL RESULT AND ANALYSIS

A. ORB features using 600 images

The histograms are plotted depending upon neighboring pixel values. Those variations of histograms are connected and plotted to get the features. Finally, using the support vector machine (SVM) classifier, the image is classified as real or fake. Experimental results show that this method having accuracy rate reaching up to 91 % with multi-resolution WLD descriptor of the images on the chrominance space, in addition to giving better discrimination than single resolution, better edge detection, and its being robust to noise change and illumination.

Classifier	Accuracy (ORB)	Precision(ORB)	Recall(ORB)
SVM+RBF	90.24	87	83
SVM+EM	97	82	87

Table 1: 600 Images +ORB Features

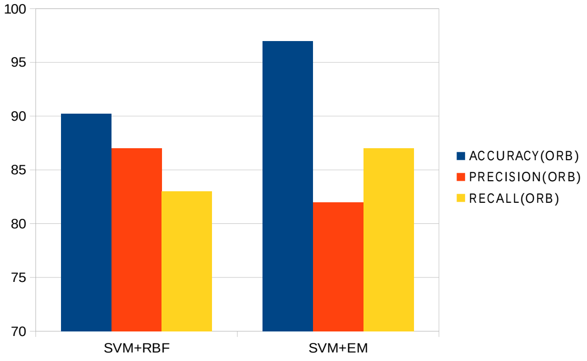

Fig.1: ORB features with 600 images

B. SIFT features using 600 images

In above figure use 600 images of COMFOD dataset which train by SVM with RBF kernel and SVM optimize by EM classifier SVM with EM classifier give more accuracy than SVM with RBF. If classification use less train images.

Classifier	Accuracy(ORB)	Precision	Recall(ORB)
SVM+RBF	90.24	87	83
SVM+EM	97	82	87

Table 2: 600 Images +SIFT Features

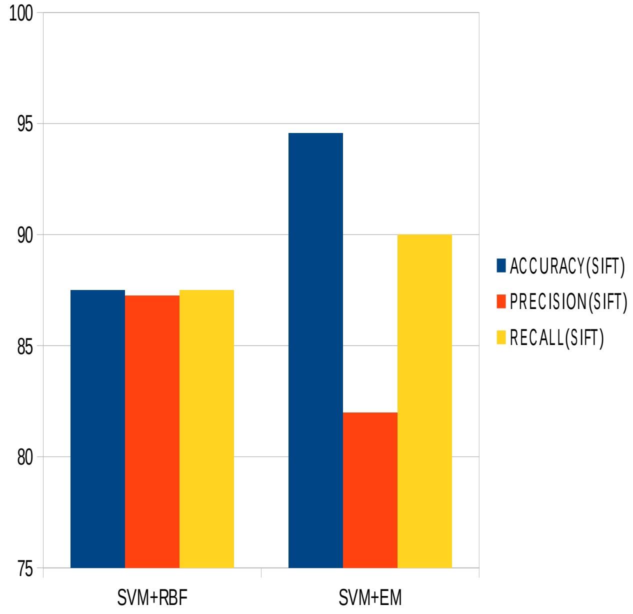

Fig.2: SIFT features with 600 images

5.3 OR B features using 300 images

Classifier	Accuracy(ORB)	Precision	Recall(ORB)
SVM+RBF	93.14	85	90
SVM+EM	94	89	80.23

Table 3: ORB features with 600 images

V. CONCLUSION AND FUTURE SCOPE

With the image processing technology rapid progress, the digital image forgery detection in forensic science is an interesting research topic. We can consider the image tampering of specific type as a "copy-move forgery", which is an emerging problem in digital image forensic field. In copy-move forgery method, original digital image part is copied and on the same original image another part it is pasted for making it as a copy forged one. "Copy-Move Forgery" classification is based on SIFT and ORB Features. In this thesis, we have used different type of classifiers like SVM and EM algorithm which classifying the images in copy and original image and that gives higher accuracy and precision and recall.

On the improved method performance bases for "copy move forgery classification" in digital images, we can highly recommend extending this research in the future to:

- Deal with problems such as rotation and scales.
- Working on videos where duplicated blocks are searched which performs on multiple image frames.
- The future digital forensic direction in conjunction would be multiplex forensic tools with awareness and sensible policy and law in which that convincing digital forgeries is created.

VI. REFERNCES

[1]. A. Fridrich, et al., Detection of Copy-move Forgery in Digital Images, 2003.

[2]. Y. Huang, et al., Improved DCT-based detection of copy-move forgery in images,Forensic Science International 206 (1–3) (2011) 178–184.

[3]. A. Popescu and H. Farid, Exposing digital forgeries by detecting duplicate image regions, Dept. Computer. Sci. Dartmouth College, Tech.Rep. TR2004 515, 2004.

[4]. B. Mahdian, S. Saic, Detection of copy-move forgery using a method based on blur moment invariants, Forensic Science International 171 (2007) 180–189.

[5]. Li Jing, and Chao Shao," Image Copy-Move Forgery Detecting Based on Local Invariant Feature Journal Of Multimedia,Vol.7,No.1, February 2012.

[6]. Vincent Christlein," An Evaluation of Popular Copy-Move ForgeryDetection Approaches", IEEE Transactions On Information Forensics And Security, 2011.

[7]. S. Bayram, H.T. Sencar, N. Memon," An efficient and robust method for detecting copy-move forgery", in: IEEE International Conference on Acoustics, Speech and Signal Processing, IEEE Press, New York, 2009.

[8]. X. Pan, S. Lyu," Detecting image region duplication using SIFT features", in: IEEE International Conference on Acoustics Speech and Signal Processing (ICASSP),2010, 2010, 1706–1709.