Copy-Move Image Forgery Detection Using Hybridization of SIFT and SVM Classifier

Yash Parashar
Research Scholar, M.Tech (Computer Science& Engineering), ITM Group of Institutions, Gwalior (M.P.), India.

Shirish Mohan Dubey
Asst. Professor, Depart. Of Computer Science and Engineering, ITM Group of Institutions, Gwalior (M.P.), India.

Abstract – By means of the beginning of influential and powerful image altering softwares, image genuineness poses large obstacle for the image forensics. Persons can no more have their trust in what they see. As soon as a section of image is copied, geometrically changed and pasted on different various sites onto the very same image with the aim of cover or concealing some of the vital information, this is copy-move forgery. In the past couple of few years numerous observes for copy move forgery detection have been anticipated. This paper is using the Discrete Wavelet Transform (DWT) and Discrete Cosine Transform (DCT) alongside with Scale Invariant Feature Transform (SIFT) and Laplacian operator aimed at the detection of copy-move image forgery. Laplacian operator is using for mining feature points of image at that time; SIFT key point descriptors are extracted for the images to identify it distinguishably. After that these feature points are coordinated (matched) and classification is done with SVM classifier.

Index Terms – Tampering, Digital image forgery, SIFT, Key points, SVM.

1. INTRODUCTION

Images have been reconditioned and altered ever since the digital photography has arisen into reality. In the past span digital image statistics were deployed in numerous fields like fashion magazines, tabloids, court room videos, papers, etc. [2]. Discovery of image forgery is a challenging task due to numerous methods of snooping with the image and huge number of varied image capturing devices accessible. Amongst numerous forgery detection methods some of them are categorized into two major approaches: first is active and another is passive. Active approaches, which are also denoted as non-blind or intrusive methods, need few resources to be injected in the original image. Due to this requirement, active approaches have forced scope. Few of the examples of these methods are watermarking and the use of the digital signature of camera. Passive approaches, also mentioned as non-intrusive or blind methods, don't require any data to be injected in the digital image.

Any digital image can be forged by put on numerous modifications like adding of blurring, noise, compressing etc. Copy Move forgery is different sort of image forgery. The ease and efficiency of copy-move forgery permits it designate the most often used forgery to modify the configuration of an image [3]. Generally the imitator forges the image through a purpose of covering some object through covering it with a section that is basically copied from the very same image. For instance if there is a only flower in rose plant image, imitator copy and pastes the very same rose to the different twigs or branches to improve the attractiveness of that image. A general illustration of copy- move forgery is presented in figure 1. In the main original image, there were only three missiles. This image was altered to display four missiles thru copying-pasting one of the missiles in that same image [3].

(a) Original Image (b) Forged Image

Figure 1 Example of Copy-Move Forgery [3]

In this very paper we presented a process to sense forgery using a grouping of DCT and DWT through which feature vector are mined. After that Harris Corner Detector alongside with normalized Laplace operator is used or applied for takeout key points. Afterwards these key descriptors catch on using SIFT and features are harmonized or coincide with the maximum similarity area merging. The SVM classifier is used here for classification and for compute results.

2. LITERATURE SURVEY

There was a way to detect copy move forgery in the images introduced using SIFT and DWT [2]. Image is sub-divided using DWT into the four sub-bands, then SIFT is used on LL portion only and the descriptor vector is initiate for given main features. To check up whether the image is forged or not, a match among different descriptor vectors is made.

A blind process or method is used for the detection of copy-move image forgery via un-decimated dyadic wavelets is introduced in [4]. By the shift invariant property, the dyadic wavelet transform (DyWT) is much more suitable for data examination as compare to discrete wavelet transform (DWT). The image is sub-partitioned into HH1 and LL1 sub-bands. Then each sub bands are separated into overlapping blocks and likeness between these blocks is calculated. The key impression of this technique is that, the match among the copied and moved blocks from the LL1 sub band must be high, while that in the HH1 sub band have to be low by the noise variation in that moved block.

A methodical and comparative analysis of SIFT and its family, together with GSIFT, PCA-SIFT, SURF, CSIFT, and ASIFT is completed in [5]. The performance and execution is measured and the consumption of time is calculated in various circumstances. The main results concluded that each of the algorithms has its own benefits.

3. PROBLEM STATEMENT

In the processing of image, there is no steady organization, association or farm so that the numerous properties of images are used as detailed. For the process of image matching, the effort starts with pre-processing stage, there are much many of detail specific extraction methods are obtainable in the literature. The former algorithms suffer in most likely the following areas:

- Low accuracy
- Problem of generating key points
- High percentage of false match ratio and false rejection ratio
- Low performance time
- Problem in the discovery of matching feature points

4. PROPOSED METHOD

The suggested method is based at discovering important points in the real image and altered image. For this first we take true and forged images for the processing. Then we resize both images into 512*512 and after that extract the features of real image via discrete cosine transform and discrete wavelet transform. Afterwards applying Discrete Cosine Transform (DCT) modulated with Discrete Wavelet Transform (DWT) to diminish the feature vector element of the image so as to extract feature of image.

Formerly with the Laplacian operator we discover the 'characteristic scale' in place of each and every FP. And if there is not any characteristic scale for FP it is rejected. („charac. scale' - it's a scale where Laplacian operator receives local maximum, and also at this scale FP is a 'corner'. Characteristic scale gives to descriptor invariance to scale/zoom).

In any image there are a so many parts of interest which can be extracted to provide feature description of the image. The algorithm SIFT can be applied to find a specific object in an image that contains various dissimilar objects. SIFT procedure or algorithm delivers a set of features which do not get modified. Another distinct advantage is that the SIFT is very resilient to noise in the image [6]. Thus, extract key descriptor of both images using SIFT method then calculate distance between two feature points using standardized distance. Match feature points using maximum similarity region merging. SVM is one of the best known methods in pattern classification and image classification [9]. It is intended to distinct of a set of exercise images 2 unlike classes, (x1, y1), (x2, y2), ..., (xn, yn) where xi in Rd, d-dimensional feature space, and yi in {-1,+1}, the class label, with i=1..n [1]. Thus, classification via Support vector machine is done for classifying the database. Thus, the accuracy, precision, recall and processing time is calculated.

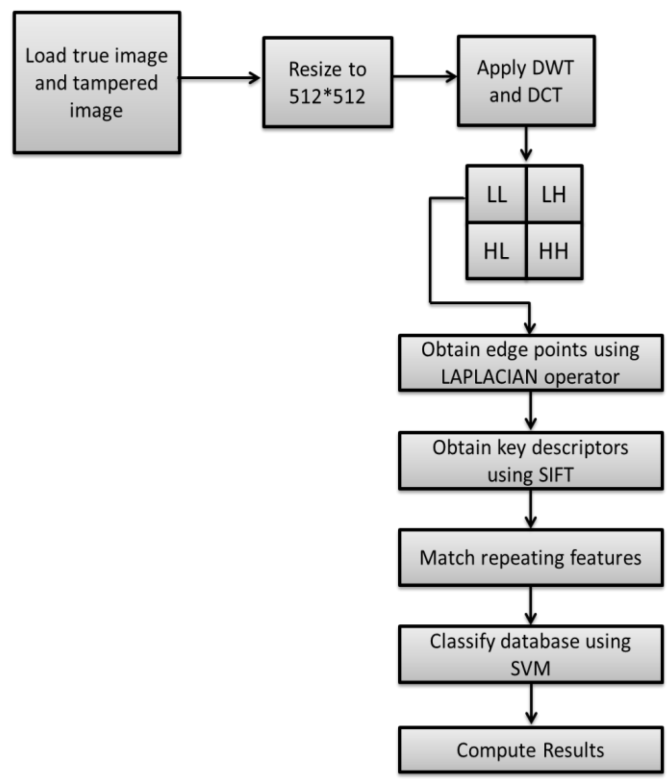


Figure 2 Flow Chart of Proposed Method

Discrete Cosine Transform

The Discrete cosine transform is a distortive compression algorithm. That is abandons individuals frequencies which do not affect or disturb the image as the humanoid eye observes it [7]. For the double dimensional DCT the mathematical function is described as follows:

$$X_{k1,k2} = \sum_{n_1=0}^{N_1-1}\left(\sum_{n_2=0}^{N_2-1} x_{n_1,n_2} \cos\left[\frac{\pi}{N_2}\left(n_2 + \frac{1}{2}\right)k_2\right]\right)\cos\left[\frac{\pi}{N_1}\left(n_1+\frac{1}{2}\right)k_1\right]$$

$$= \sum_{n_1=0}^{N_1-1}\sum_{n_2=0}^{N_2-1} x_{n_1,n_2} \cos\left[\frac{\pi}{N_1}\left(n_1+\frac{1}{2}\right)k_1\right]\cos\left[\frac{\pi}{N_2}\left(n_2 + \frac{1}{2}\right)k_2\right]$$

Discrete Wavelet Transform

The Wavelets are distinctive functions which, in a form alike to sine and cosine in the Fourier analysis are used as basal functions for representing signals. For 2-D images, enforcing DWT resembles to proceeding out the image via 2-D filters in for each dimension [7]. The filters split the input image into the 4 no overlapping multiresolution sub bands LH1, LL1, HH1 and HL1. The first sub-band LL1 signifies the coarse scale DWT coefficients although the other sub bands LH1, HL1 and HH1 signify the fine-scale of DWT coefficients. Through its excellent spatio-frequency localization attributes it permits the exploitation manipulation of the masking result of the human visual system in order that if a DWT literal-coefficient is altered, only the region or area corresponding to that literal-coefficient will be modified.

Laplacian Operator

Laplacian Operator helps in discovering edges in an image and is also referred to as derivative operator. Prewitt, Sobel, Robinson and Kirsch operators differ from Laplacian Operator as all are first order derivative masks whereas Laplacian operator is a second order derivative mask.

The Laplacian L(x,y) of an image by pixel intensity values I(x,y) is given by:

$$\mathrm{L(x.y)} = \frac{\partial^2 I}{\partial x^2} + \frac{\partial^2 I}{\partial y^2}$$

Scale Invariant Feature Transform (SIFT)

The SIFT method is applied to abstract distinctive attributes from an image. A very important advantage of this transform is that it provides a set of features which are invariant to scaling and rotation [2]. The scale gap or space of any image is showed as the function L(x, y, σ), that is formed by convolving G(x, y, σ) and that's variable scale Gaussian by an input image I(x, y). For Key-point descriptors gradient orientation histogram is used, this will provide robust representation. First compute comparative location and magnitude in 16*16 neighborhoods at key points, divide the 16*16 region into 4*4 blocks.

Support Vector Machine (SVM)

Classification is a procedure of allocating an unknown data sample from among the pre-defined classes. SVM is a classifier that is modelled using data with known classes [9]. This procedure contains of two stages: training phase and testing phase. In the training phase, the system is learning to find a mapping (classifier or model) between the features extracted from the images in the training set and their classes while in the testing phase, the system uses the learned model and the features extracted from new images (i.e. testing set) to assign them to classes.

5. RESULTS AND DISCUSSIONS

The figure (3), (4) and (5) explains the various steps of the proposed method. The algorithm was tested and satisfied using a set of 30 images out of which 15 were original and 15 were forged images in the database.

Figure 3 (a) True Image (b) Tampered Image

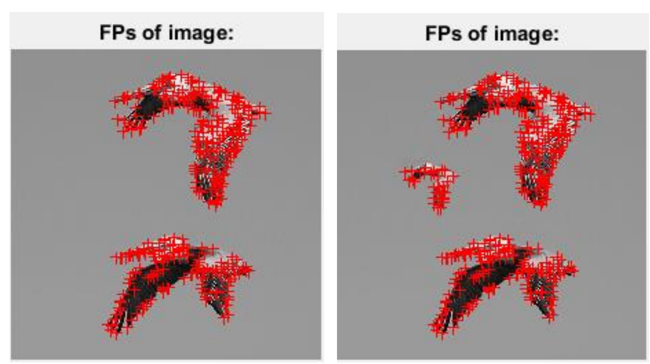


Figure 4 (a) True Image FPs (b) Tampered Image FPs

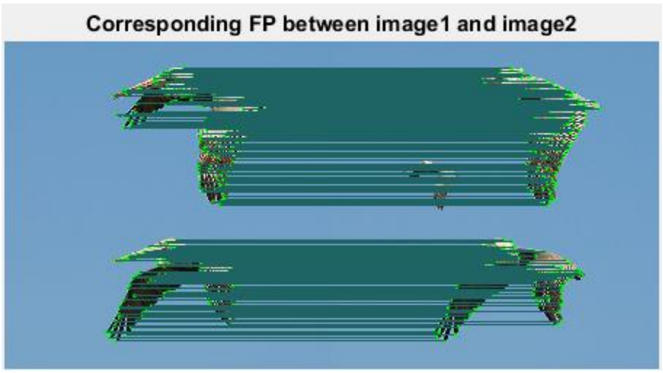


Figure 5 Similar Feature Points Matching Between True Image and Forged Image

The performance of any system can be measured in terms of sensitivity, specificity and accuracy. Sensitivity relates to the ability of the algorithm to detect a forged image correctly as forged. Specificity relates to the ability of the algorithm to identify an authentic image correctly as authentic. Hence a high value of sensitivity and specificity imply better performance of the system.

$$sensitivity = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}}$$

$$specificity = \frac{\mathrm{TN}}{\mathrm{TN} + \mathrm{FP}}$$

$$accuracy = \frac{\mathrm{TP} + \mathrm{TN}}{\mathrm{TP} + \mathrm{FP} + \mathrm{FN} + \mathrm{TN}}$$

where terms TP, FN, TN, FN are,

TP (True Positive): Forged image identified as forged

FP (False Positive): Authentic image identifies as forged

TN (True Negative): Authentic image identified as authentic

FN (False Negative): Forged image identified as authentic

F P R = 1- specificity (FPR: False Positive Rate)

FNR = 1- sensitivity (FNR: False Negative Rate)

Algorithm	**Accuracy (%)**	**FPR (%)**	**FNR (%)**	**Processing Time (in sec.)**
Reference method	55.56%	4%	8%	11.3011
Proposed method	83.90%	0.6%	0.25%	0.0505

Table 1: Results of the Proposed Method

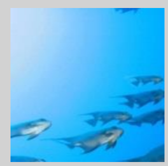
True Images	**Forged Images**	**Reference Processing Time**	**Proposed Processing Time**
		5.0889	0. 0242
		7.1283	1.6296
		3.2668	0.0242

Table 2: Performance Results of Image

6. CONCLUSION AND FUTURE WORK

Various methods of image forgery have been developed and are being used at present. Former algorithms for copy-paste forgery detection were studied. The proposed method uses combination of both DCT and DWT for feature extraction. Then, Laplacian operator along with SIFT is applied. Henceforth the proposed algorithm identifies image forgery although the copied part is modified and pasted. The results were computed on the database of 30 images (15 authentic and 15 forged images). The method which is proposed gives rational values of sensitivity and specificity. Resilience of the suggested algorithm was tested by checking tampered images where the copied part is modified and pasted. Then a comparison is made between proposed method and the existing method and it was found that the proposed method gave higher accuracy than the existing method. The time required for computation was less and therefore it can be inferred that the computational complexity was reduced. In further course of work, we would like to work on rotation attack with the existing method.

REFERENCES

[1] Digital Image Processing by Rafael C. Gonzalez and Richard E. Woods.
[2] Ms.P.G.Gomase, Ms. N.R. Wankhade, "Advanced Digital Image Forgery Detection: A Review", International Conference on Advances in Engineering & Technology-2014, e-ISSN: 2278-0661, p-ISSN: 2278-8727 PP 80-83.

[3] Mohammad Farukh Hashmi, Aaditya R. Hambarde, Avinash G. Keskar, "Copy Move Forgery Detection using DWT and SIFT Features", 978-1-4799-3516-1113, IEEE, 2013.
[4] Vidhi P.Raval, "Analysis and Detection of Image Forgery Methodologies", International Journal of Scientific Research and Development 2013, Vol. 1, Issue 9, 2013.
[5] Ghulam Muhammad, Muhammad Hussain , George Bebis, "Passive Copy Move Image Forgery Detection using Undecimated Dyadic Wavelet Transform", Elsevier, 2012.
[6] Jian Wu, Zhiming Cui, Victor S.Sheng, Pengpeng Zhao, Dongliang Su, Shengrong Gong, "A Comparative Study of SIFT and its Variants", Measurement Science Review, Volume 13, No.3, 2013.
[7] Nehaverma, Deepak Sharma "Advanced Maximal Similarity Based Region Merging By User Interactions", IJERA Volume 3, Issue 4, pp.681-688, Jul-Aug 2013.
[8] Jian Wu, Zhiming Cui, Victor S.Sheng, Pengpeng Zhao, Dongliang Su, Shengrong Gong, "A Comparative Study of SIFT and its Variants", Measurement Science Review, Volume 13, No.3, 2013.
[9] Vandana Bavkar "Digital Watermarking Using Combined DWT And DCT", International Journal of Engineering Research & Technology (IJERT) Volume 2, Issue 4, April 2013.
[10] Le Hoang Thai, Tran Son Hai, Nguyen Thanh Thuy "Image Classification using Support Vector", I.J. Information Technology and Computer Science, 32-38, Machine and Artificial Neural Network(MECS), May 2012.